2017

ADAPTIVE MULTI-OBJECTIVE OPERATING ROOM PLANNING WITH STOCHASTIC DEMAND AND CASE TIMES

Vivek Reddy Gunna
University of Kentucky, vivekreddy.gunna@uky.edu
Author ORCID Identifier:
https://orcid.org/0000-0003-4519-928X
Digital Object Identifier: https://doi.org/10.13023/ETD.2017.494

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Vivek Reddy Gunna, Student

Dr. Wei Li, Major Professor

Dr. Haluk E. Karaca, Director of Graduate Studies

ADAPTIVE MULTI-OBJECTIVE OPERATING ROOM PLANNING WITH STOCHASTIC DEMAND AND CASE TIMES

THESIS

A thesis submitted in partial fulfillment of the
requirement for the degree of Master of Science in Mechanical Engineering
in the College of Engineering
at the University of Kentucky

By

Vivek Reddy Gunna

Lexington, Kentucky

Co-Directors: Dr. Wei Li, Assistant Professor of Mechanical Engineering

and Dr. Fazleena Badurdeen, Associate Professor of Mechanical Engineering

Lexington, Kentucky
2017

ABSTRACT OF THESIS

ADAPTIVE MULTI-OBJECTIVE OPERATING ROOM PLANNING WITH STOCHASTIC DEMAND AND CASE TIMES

The operating room (OR) is accountable for most hospital admissions and is one of the most cost and work intensive areas in the hospital. From recent trends, we discover an unexpected parallel increase in expenditure and waiting time. Therefore, improving OR planning has become obligatory, particularly regarding utilization, and service level. Significant challenges in OR planning are the high variations in demand, processing times of surgical specialties, the trade-off between the objectives, and control of OR performance in long-term. Our model provides OR configurations at a strategical level of OR planning to minimize the tradeoff between the utilization and service level accounting for variation in both demand and processing times of surgical specialties. An adaptive control scheme is proposed to aid OR managers to maintain the OR performance within the prescribed controllable limits. Our model is validated using a simulation of demand and processing time data of surgical services at University of Kentucky Health Care.

Keywords: Operating Room, utilization, service level, and Trade-off.

Vivek Reddy Gunna

Student's Signature

12/12/2017

Date

ADAPTIVE MULTI-OBJECTIVE OPERATING ROOM PLANNING WITH STOCHASTIC DEMAND AND CASE TIMES

By

Vivek Reddy Gunna

Dr. Wei Li

Co-Director of Thesis

Dr. Fazleena Badurdeen

Co-Director of Thesis

Dr. Haluk E. Karaca

Director of Graduate Studies

12/12/2017

Date

ACKNOWLEDGEMENTS

I want to thank Dr. Wei (Mike) Li for his guidance and encouragement throughout this research. His focus and critical thinking on the research has inspired me to accomplish this task. I would like to appreciate his persistence in having research meeting twice in a week, defining every problem in PORM scheme: Problem, objective, results and Methodology improved my presentation skills.

I would like to thank Dr. Jawahir and Dr. Badurdeen for agreeing to be a part of my committee. I sincerely believe, their constructive criticism and comments on this report has improved the quality immensely.

I also would like to thank my parents and friends for their and support and belief in me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1 Introduction

1.1. Background

Operating room (OR) is essential because of rising demand, increasing health care costs (Figure1. 1) and waiting lists. U. S. healthcare expenditure was increased by $ 0.2 trillion in years 2014-15 (Forbes,2015) and waiting time is increasing every year (Viberg et al., 2013). Ironically, despite increasing expenditure, hospitals are unable to reduce the waiting. This inadequacy is attributed to the inefficient OR planning in the hospitals. The OR is one of the most cost and work intensive areas of a hospital. The OR's are the primary reason for almost 70% of all hospital admissions (Ehrenfeld et al., 2013) and account for more than 40% of a hospital's total revenue. Therefore, OR managers are consistently looking for ways to maximize the utilization, service level, patient flow, and minimize the waiting time and cost.

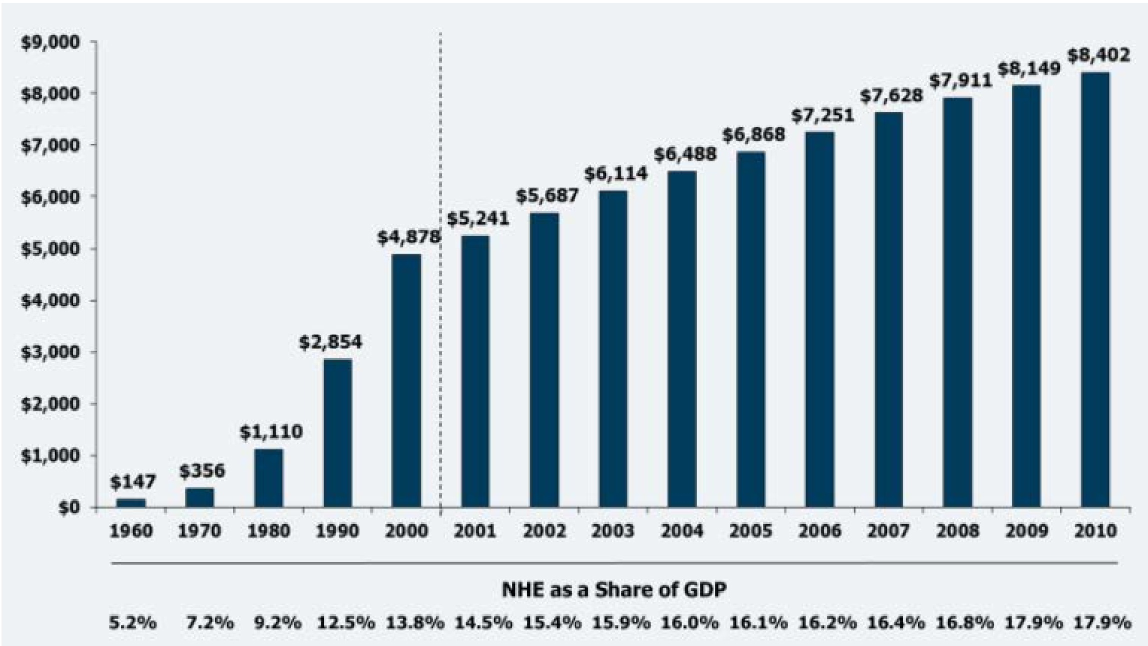

Figure 1. 1:National Healthcare Expenditure (NHE) per capita

1.2. Motivation

OR planning has always been very complicated because of the phases of planning, sequential stages, highly stochastic demand, processing times, the sheer diversity of the surgical services and the priorities of the stakeholders, patients and OR managers. OR planning is carried out in three hierarchical phases (Vissers et al., 2001) strategic, tactical and operational respectively. The strategic plan is carried out for a long-term where, agreements with surgical specialties concerning their patient volumes, targets, etc. are set up. Tactical level planning addresses the usage of resources on a medium-term by developing cyclic master scheduling strategy (MSS). Operational phase deals with resource re-allocation and re-sequencing resulting from dynamic disturbances in healthcare systems, such as variations in processing times, and fluctuations in demand, e.g., no-shows, cancellations, and emergencies (Banditori et al., 2013).

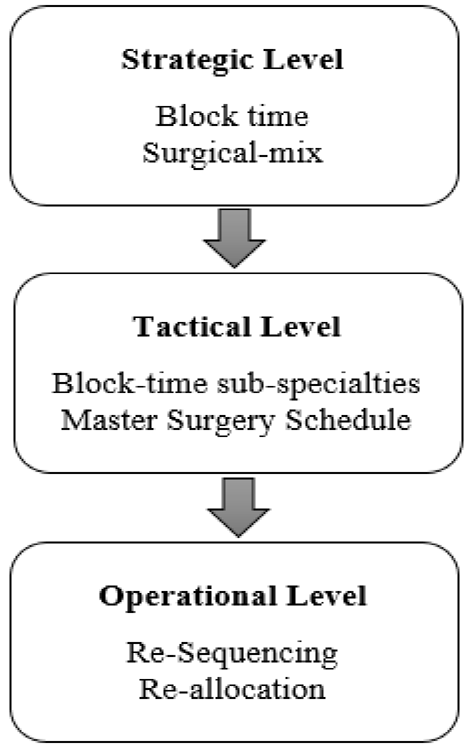

Figure 1. 2:Planning Phases

Strategic OR planning phase is significant, because of its impact on other planning phases down the line. Strategic OR planning deals with the allocation of block time: allocated time for each surgical specialty, and surgical-mix: planned number of cases to be served for each specialty. Both the block time and surgical-mix together are referred as configuration. It is imperative to establish one.to.one relationship between the configurations and Key performance indicators (KPI's) of OR performance, to maintain the performance of OR within controllable limits given the high variation in the system.

Apart from the planning, OR performance is also affected by the upstream and downstream stages of the peri-operative process. Perioperative process deals with the surgical interventions at the hospital. The Peri-operative process can be broadly broken down into three stages as shown in Figure 1.1 (Gupta, 2007): pre-operative stage, intra-operative stage, and post-operative stage. Pre-operative stage deals with the preparation of the patient for the surgery with counseling, anesthesia, etc. Intra-operative stage constitutes the OR, where the actual surgery is performed. Depending on the condition of the patient after the surgery, the patient is moved either to post-anesthesia care units (PACU) or intensive care unit (ICU) in post-operative stages of recovery.

Figure 1. 3:Peri-operative process

Delay at the pre-operative stage leaves the OR idle leading to low utilization and may cause overtime when other case is scheduled into that OR (Roberts et al., 2015). While non-availability of beds in post-operative stages leads to blocking, in which case the patient must recover in OR itself, blocking the sequence of surgeries leading to high waiting time and increased cost (Augusto et al., 2010), (Wang et al., 2015). Abedini et al., (2017) proposed an optimization model along the peri-operative process to reduce the blocking of cases among pre-operative, intra-operative and post-operative stage. Given the high cost associated with OR relative to other stages, hospitals usually prefer standard OR plans and more resources in pre-operative and post-operative stages concerning Intra-operative stage to smoothen the patient flow. However, additional disturbances in a system like the variation in demand and processing times, coupled with disruptions from upstream and downstream stages, emergencies, cancellation affect the KPI's of OR.

KPI's of OR performance from the literature are (Cardoen et al., 2010) utilization, service level, waiting time, overtime, idletime, revenue per OR minute, cost and patient flow. Significant KPI's among them are, utilization and service level, as they have direct relationships with other KPI's. For example, maximizing utilization reduces the waiting time, idletime, and cost, and maximizing service level can be credited with increasing the patient flow and revenue per OR minute. Utilization is the ratio of the used time to the allocated time. Service level is the ratio of the actual demand to the planned number of cases. Therefore, research objective is to develop a standard strategic OR plan which maximizes the utilization and service level given the variation in demand and processing times.

1.3. Challenges in OR planning

One of the primary problems in developing the efficient OR plans is the stochastic demands and processing times of various surgical specialties. Unlike production facility where the uncertainty in demand and processing time are relatively low, efficient production plans can be carried out with little or no disruption in the program. Nevertheless, due to the high fluctuation in the OR, OR planning requires an accommodating plan which dampens the high variation in demand and processing times and simultaneously balances the overtime and idletime.

Another significant challenge arises from the objectives of the two principal stakeholders of the OR department, OR managers and patients. Hospital management usually aspires to efficiently utilize its resources, which increases their revenues and cuts cost. On the other hand, patients prefer high service rate, short waiting time and low cost. These Objectives are inconsistent with each other, because reserving too much OR time may improve the service rate (reduces waiting time), but performs low on utilization generating more idletime, and consequently high cost. On the other hand, packing limited block time with more number of cases to maximize utilization causes higher overtime. Therefore, the objective from the operations research perspective is to develop OR plans which minimize the cost by balancing overtime and idletime and minimizes the tradeoff between utilization and service rate.

Although Optimal OR plans are generated to maximize the utilization and service level, there is a variation in these KPI's in real-time due to the continuous fluctuation in demand and processing times of surgical services. This phenomenon presents a challenge to maintain the OR utilization and service level in control in the long term. Therefore, it is of great importance to develop a dynamic adaptive OR planning scheme, to change the OR configuration according to the current performance, to maintain the KPI's within a controllable limit in the long term.

1.4. Contribution

Contributions of our work in OR planning are: (1) Prove that existence of trade-off between utilization and service level; (2) Multiple portfolio optimization to minimize the trade-off between utilization and service level; (3) An adaptive control scheme to maintain utilization and service levels within the controllable limits in the long term. (4) Validating the OR configurations and an adaptive scheme using the historical distribution of demand and processing times of surgical services at University of Kentucky Healthcare (UKHC) using a simulation along the time horizon.

First, we balance the cost incurred due the overtime and idletime generated due to the variation in demand and processing time using the newsvendor model as proposed in

Strum et al. 1997. We present the trade-off between utilization and service level by modeling the configurations from two perspectives: (1) demand perspective: to maximize the service level (2) workload perspective: to maximize the OR utilization.

Second, using historical data of utilization and service level, we provide an efficient frontier of configurations which minimize the trade-off between the utilization and service level using multiple objective portfolio optimization, with different preferences among the objectives. This optimization provides one to one relationship between the configurations and the expected performance of OR regarding utilization and service level.

Third, we developed an adaptive control scheme which monitors the error in utilization and service level from the targets for the current time and changes the OR configuration adaptively to maintain the utilization and service level within the predetermined controllable limits by the OR manager.

Fourth, we validate the performance of our model using statistical process control (SPC) and control charts by simulating normally distributed demand and processing times of major surgical services provided at the UKHC.

1.4. Impact Statement

Performance of Hospitals are judged based on important KPI's like utilization, cost, waiting time, throughput time, service level etc... OR mangers and patients are two important stakeholders of the OR. OR managers often strive for an efficient OR planning schemes to maximize utilization which reduces cost and maximizes service level to reduce waiting time for patients. OR plans are often disrupted by the stochastic demands and case times leading to long waiting lists and high costs for patients. This research will aid in realizing efficient OR planning in hospitals to reduce the cost of healthcare and waiting time by increasing the utilization of resources and service level given the stochastic demands and case times. Adaptive control scheme is also illustrated to aid OR managers to maintain the OR performance measures within prescribed control limits.

1.5. Thesis structure

The rest of the thesis is organized as follows:

Chapter 2 provides a literature review, on the application of operations research in healthcare, the status of multi-objective optimization in strategic OR planning, state of literature dealing with the variation in demand and processing times, application of newsvendor model and portfolio optimization in the healthcare background.

Chapter 3 presents the methodology. Firstly, a detailed problem description of OR planning and evaluation schemes of the OR performances is provided. Secondly, the trade-off among the utilization and service level is staged by modeling the configurations from demand and workload perspectives. Thirdly, based on the historical data of utilization and service level, optimal OR configurations are formulated using a multiple-objective portfolio optimization which minimizes the trade-off between the utilization and service level. Fourthly, a detailed description of an adaptive control scheme is given, which ensures that the utilization and service level is within the controllable limits along the time horizon.

Chapter 4 provides the results of the case study. A case study is carried out with the historical data of surgical services at UKHC. First, utilization and service level are compared among two sets of OR configuration each developed from demand and workload perspective, to show the trade-off among these KPI's of OR. Second, efficient portfolio frontiers to minimize the trade-off between utilization and service level are generated. Third, the adaptive control scheme is validated by verifying the conformance of utilization and service levels within the controllable limits along the time horizon.

Chapter 5 discusses conclusions and directions for future research.

Chapter 2 Literature review

This chapter emphasizes on the literature review in five specific topics. First, general trends regarding the application of operations research in OR management are introduced. Second, literature dealing with significant objectives in OR planning is discussed in detailed. Third, literature studying the impact of variation in demand and processing time on KPI's of OR are elaborated. Finally, optimization techniques used in this research, newsvendor model, and portfolio selection are introduced, and their applications in OR planning are reviewed.

2.1 Operations Research in OR management

OR management is an extensive and complex field of study because of the hierarchical planning structure, stochastic demand, processing times, multiple objectives to accomplish and the trade-offs among the objectives. Guerriero and Guido (2011) presented a structural literature review on how Operational Research can be applied to the surgical planning and scheduling processes. Cardoen et al., (2010) summarizes the significant trends in research on operating room planning and scheduling and identified areas to be addressed in the future. Erdogan and Denton (2011) presented a thorough literature review on, challenges and directions for future research in OR planning and scheduling from operations research perspective.

2.2 Objectives in OR planning

OR planning is carried out in three hierarchical phases (Vissers et al., 2001) strategic, tactical and operational respectively. We focus on strategic phase planning, which requires a decision on the block times and surgical-mixes for surgical specialties for a long term. As the decisions made in the strategic phase directly impact the following two phases, it is imperative for OR managers to address multiple objectives in strategic OR planning. However, Majority of the current literature focuses on single objectives like maximizing utilization, profit, patient flow and minimize waiting time. The following sections will elaborate on the research specific to utilization, service level, and multiple objectives.

2.2.1 Utilization

Utilization is a thoroughly studied objective in the literature and is one of the critical KPI's of OR performance. Ozkarahan (2000) used a goal programming approach to schedule cases into OR to maximize utilization, under constraints like surgeon preferences, intensive care capabilities, and available time restrictions. However, this model has strong assumptions of accurate estimation of surgeon-specific surgical duration and availability inventory of case, which does not hold true in the actual OR setting. Dexter and Traub (2002) provided a heuristic to schedule elective cases into OR's using the sample mean from the historical data to maximize OR utilization.

Kharraja et al., (2006) proposed a master surgical schedule approach to maximize utilization. A cyclic master surgery schedule is developed for a week using integer linear programming first. Then, they introduced multiple knapsack problem to assign additional to exploit the unused OR time generated because of the variation in processing times and cancellations. Ye et al., (2017) proposed an efficient sequencing heuristic to minimize the total completion time and associated it to maximizing utilization. Dexter et al., (2005) discussed that OR utilization is highly unstable in the presence of high variations in processing times and demand. Improving utilization needs OR managers to reserve adequate block time to avoid both idletime and overtime. In other words, balance the tradeoff between reserving too much time leading to low utilization and too little time resulting in higher overtime.

2.2.2 Service level

Service level has manifold of definitions in the literature. Service level is addressed as throughput: number of treated patients in a period, also referred as patient flow. Baligh and Laughhunn (1969) proposed a linear model for resource allocation to maximize patient flow under constraints like resources, no. of a patient available, budget, and policy constraints. VanBerkel and Blake (2007) studied the impact redistribution of capacity among surgical specialties according to the variation in demand using simulation. This research provided multiple options in capacity planning to decrease the waiting time for elective surgeries. Santibáñez et al., (2007) developed a mixed integer linear programming model to schedule patients into OR's, and reported an increase in the number of cases served with same the capacity. Testi et al., (2007) proposed using bin packing algorithm to generate master schedule strategy (MSS) which maximizes throughput with deterministic processing times. Abedini et al., (2016) optimized operating room planning by assigning priorities among surgical-mixes. The above literature explicitly did not address the variation in case times and demands which have a significant effect on OR performance.

The service level of an OR configuration depends on the surgical-mix. Adan and Vissers (2003) proposed an integer linear programming model to optimize the surgical-mix given the target of the length of stay, and utilization of the resources. Finding the surgical-mix has been studied by researchers, Wagner (1969), discussed the possibility of formulating the Hospital Surgical-mix Selection Problem (HCMSP) as a product mix problem. Blake and Carter (2002), proposed a goal programming approach to solve the HCMSP from cost and volume perspectives at the planning phase.

2.2.3 Multiple objectives

A vital aspect of OR planning is addressing the multiple objectives like utilization, service level, revenue, waiting time, etc. Reddy Gunna et al., (2017) proposed using portfolio optimization technique to model OR configurations which maximize patient flow and benefit for the hospital. Mulholland et al., (2005) employed linear programming to maximize the financial outcomes to hospitals and physicians. Zhang et al., (2009) proposed a method of allocating operating room capacity to specialties to maximize the patient flow and minimize the cost using mixed integer programming.

2.3 Newsvendor model

Newsvendor model (Porteus, 2002) is a mathematical model, used to determine optimal inventory levels, subjected to fixed cost ratios (i.e., C_o: overage cost and C_u: underage cost, $C_o, C_u > 0$) and uncertain demand $D \sim \mathcal{N}(\mu_D, \sigma_D^2)$. Newsvendor model the trade-off between overage and underage cost and minimizes the overall total cost. Overage cost is the holding cost, occurred when the actual demand is greater than the inventory level. underage cost is the setup cost or the lost sales, when the actual demand is greater than the inventory level. A simple example is illustrated in Figure 2.1, newsvendor model points the equilibrium point of holding cost and set-up cost, at which the total cost is minimum.

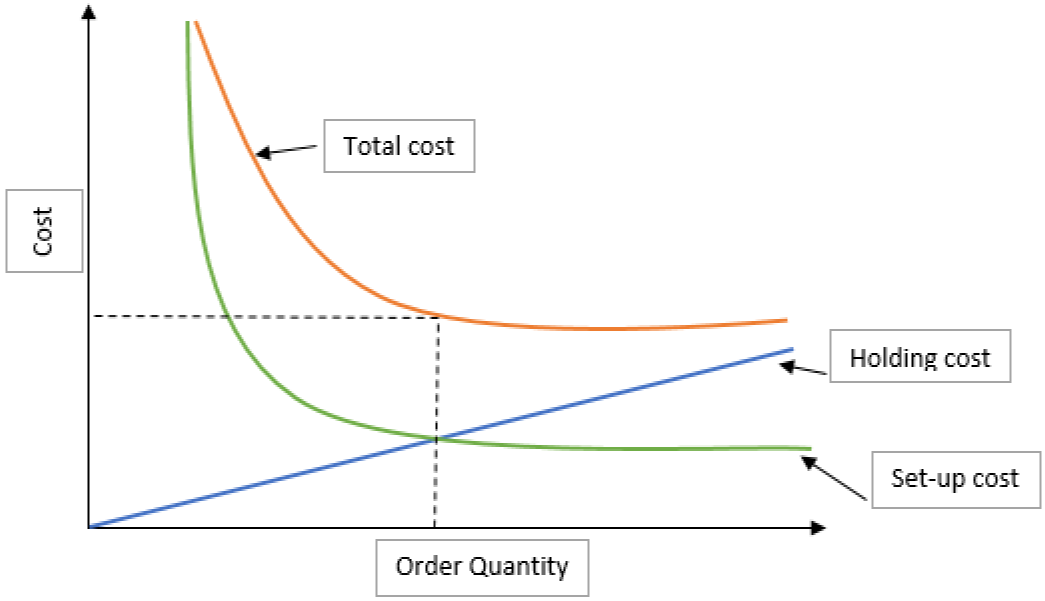

Figure 2. 1:Newsvendor model

The optimal inventory level Q^* is given by $Q^* = \mu_D + z\sigma_D$ where, $F(Q^*) = \Phi(z) = \frac{C_u}{C_o+C_u}$, and $z = \frac{Q^*-\mu_D}{\sigma_D}$. Strum et al., (1997) modeled the tradeoff among overtime and idletime in block times as a newsvendor problem. Newsvendor model is extended to OR planning. Block time to allocated is regarded the order quantity, and overtime and idletime are regarded as the holding cost and setup cost respectively. Therefore, an optimal Block time from the newsvendor approach is the block time, balancing both the overtime and idletime costs. Olivares et al., (2008) extended the application of news vendor to healthcare by structural estimation framework to show the tradeoff between the overtime costs and idletime costs. However, newsvendor model does little to address the variation in demand directly, which might cause unreasonably long waiting lists for some surgeries. Variation in demand directly impacts the service level and waiting time. Therefore it is crucial to estimate the optimal surgical-mix to reduce long waiting lists. Thus, both the surgical-mix and block time should be optimized at the strategic phase to generate an efficient MSS at the tactical level.

2.4 Portfolio selection

Portfolio selection (PS) is the process of choosing a portfolio of securities by gauging various portfolios with different weighting for stocks regarding risk and reward by evaluating the historical performance (Markowitz, 1952). The objective of portfolio selection is to invest x_i proportion of total investment in n securities with average return $\overline{r}_i$ to maximize the expected reward $E = \sum_{i=1}^{n} \overline{r}_i x_i$. The constraints in portfolio selection are the sum of investment proportions is equal to one $\sum_{i=1}^{n} x_i = 1$ and the risk σ is less than the prescribed limit. The risk of a portfolio is defined as the standard deviation of expected return given by $\sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j}$ where σ_{ij} is the covariance of i^{th} and j^{th} securities.

The portfolio which provides the maximum possible expected return $\overline{E}$ can be derived from the mathematical formulation given Eqn. (2.1- 2.2). The objective function Expected Return Eqn. (2.1) is maximized subject to the constraint Eqn. (2.2), sum of all weightage is equal to one. The associated risk $\overline{\sigma}$ with the portfolio is evaluated by $\sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j}$.

$$\text{Maximize: } \overline{E} = \sum_{i=1}^{n} \overline{r}_i x_i \tag{2.1}$$

Subject to:

$$\sum_{i=1}^{n} x_i = 1 \tag{2.2}$$

The portfolio which gives the minimize possible Risk $\underline{\sigma}$ can be derived from the mathematical formulation given in Eqn. (2.3- 2.4). The objective function Risk Eqn. (2.1) is minimized subject to the constraint Eqn. (2.2), sum of all weightage is equal to one. The associated Expected return $\underline{E}$ with the portfolio is evaluated by $\sum_{i=1}^{n} \overline{r}_i x_i$.

$$\text{Minimize:} \underline{\sigma} = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j} \tag{2.3}$$

Subject to:

$$\sum_{i=1}^{n} x_i = 1 \tag{2.4}$$

Intermediate portfolios between the extremes are derived by using incremented value of risk as a constraint. The mathematical formulation to derive the intermediate portfolios is given by Eqn. (2.5-2.7).

$$\text{Maximize:} E = \sum_{i=1}^{n} \overline{r}_i x_i \tag{2.5}$$

Subject to:

$$\sum_{i=1}^{n} x_i = 1 \tag{2.6}$$

$$\sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j} \leq \underline{\sigma} + i \tag{2.7}$$

The objective is to maximize the expected return Eqn. (2.5) Subject to the constraints Eqn. (2.6), sum of all then weight is equal to one. Constraint Eqn. (2.7), the associated risk is less than or equal to the sum of the minimum risk and increment value.

An example of portfolio optimization with three assets in illustrated in Figure 2.1.

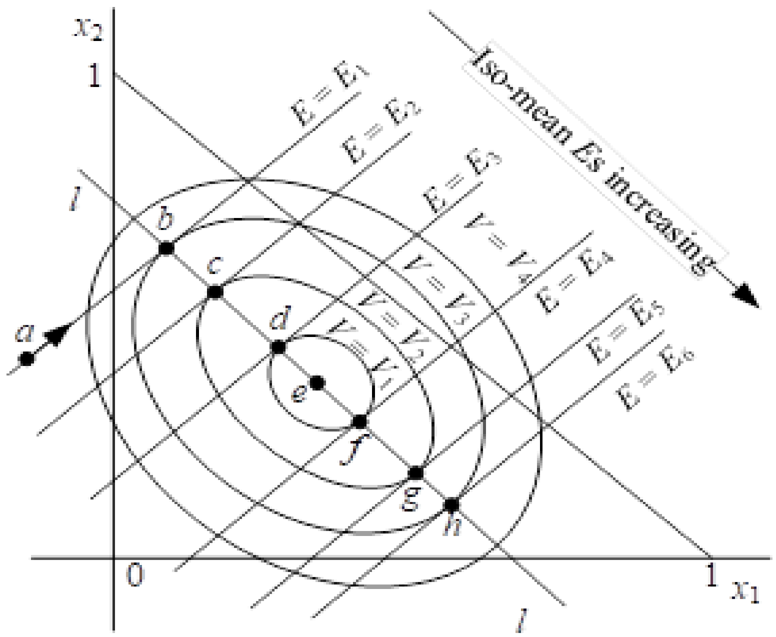

Figure 2. 2:Iso-mean and Iso-variance lines

A two-dimensional graph is used represent a portfolio with weightages x_1 and x_2 on x-axis and y-axis respectively. Since the sum of weightages in portfolio optimization is equal to one. The weightage of the third asset x_3 is given by subtracting the sum of x_1 and x_2 from one. Figure.2.1 presents the Iso-mean $\{E_1, E_2 \cdots E_6\}$ and Iso-variance lines $\{V_1, V_2 \cdots V_4\}$ of portfolios. Iso-mean lines is the locus of portfolios which have equal expected reward. Iso-variance ellipse is the locus of portfolios which have equal risk, i.e. standard deviation of expected return. The objective of the portfolio optimization is choosing the portfolios which offer maximum expected reward given fixed risk limit. From figure 2.1 we can observe that, though portfolios (points in graph) *c* and *g* have same risk, because they lie on same Iso-variance ellipse, *g* has more reward than *c*. similarly points *f* and *h* dominate *b* and *d* respectively. Therefore, it can be inferred that the efficient frontier constitutes the line formed by the points *e, f, g, h*. However, when the number of assets is greater than three, it is not possible to present the frontiers with weightages x_g. Efficient frontier in such cases is presented as a relationship between risk and reward as shown in Figure 2.2. The horizontal axis represents the risk, which the standard deviation of expected reward. The vertical axis represents the reward, which is the expected return. A portfolio is on the efficient frontier, meaning there is no other portfolio which can deliver greater rewards without increase the risk.

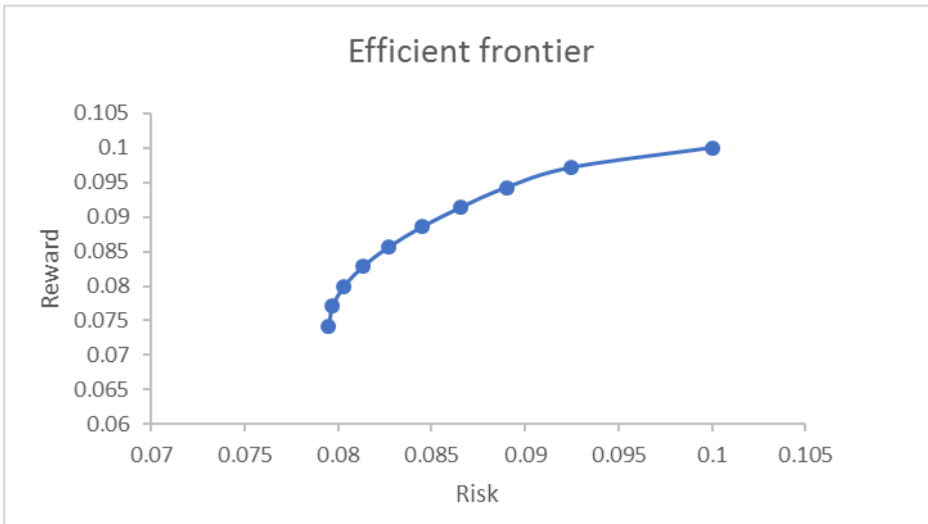

Figure 2. 3:Efficient frontier of portfolios

As far as the application of PS in healthcare is concerend, Van Houdenhoven et al., (2007) used portfolio effect and mathematical algorithms to bring down the total required OR times. Hans et al., (2008), proposed a concept of planned slack time to maximize utilization and minimize the risk of overtime, under the presence of variations in processing times at the tactical phase. A large inventory of elective cases is assumed, thus avoiding the variation in demand. Dexter and Ledolter (2003) addressed the problem of OR capacity expansion using mean-variance analysis of a portfolio of surgeons, as per their contribution margin per OR hour.

However, traditional PS generates an efficient frontier of portfolios with only one objective, in this case, it can be utilization or service rate. Determining an effective portfolio of block times and surgical-mix concurrently can be considered as a multiple portfolio optimization problem. Wang, (1999) introduced a routine to resolve the multiple benchmark portfolio optimization problem.

Existing literature does not address the variations in processing times and demands simultaneously. Moreover, the focus is on single objectives like maximizing utilization, profit, minimizing waiting time, etc. We offer a theoretical account for efficient allocation of resources among surgical specialties, block times and surgical-mixes, analyzing the expected service level, utilization, and their variation. We demonstrate that our approach is empirical; can be applied at any level of planning. These configurations minimize the tradeoff between utilization and service rate. Furthermore, our model also accounts for the stochastic nature of the surgical process, stochastic demand and processing times.

Chapter 3 Methodology

Chapter 3 presents the methodology. First, a detailed problem description of OR planning and evaluation schemes of the OR performances is provided. Second, the trade-off among the utilization and service level is staged by modeling the configurations from demand and workload perspectives. Third, based on the historical data on utilization and service level, optimal OR configurations are formulated using a multiple-objective portfolio optimization which minimizes the trade-off between the utilization and service level. Finally, a detailed description of an adaptive control scheme is given, which ensures that the utilization and service level are within the controllable limits along the time horizon.

3.1 Problem description

OR planning deals with the decision making regarding the OR configuration, which constitutes deciding on the surgical-mix: Planned number of patients of a specialty g to be treated N_g, and the Block time: Reserved time allocated for each specialty B_g. These decisions are based on the historical distribution of the demand for each specialty $d_g \sim \mathcal{N}(\mu_g^d, {\sigma_g^d}^2)$ and the distribution of processing times of each specialty $p_g \sim \mathcal{N}(\mu_g^p, {\sigma_g^p}^2)$.

Assumptions in OR planning:

1. Demand and processing times of specialties follow a normal distribution.
2. All the other resources like surgeons, anesthesiologists, nurses, equipment, etc. are available, after a standard OR plan is established.

3. All the Demand of a specialty is treated, which is less than or equal to the planned number of cases ($d_g \leq N_g$) irrespective of the overtime or idletime.

OR performance is evaluated interms of utilization and service level. Utilization and service level can be calculated for individual specialties and for the overall OR configuration. Utilization of a specialty is defined as the ratio of workload: the sum of processing times of cases treated, to the allocated block time. Utilization for each specialty is calculated using Eqn. (3.1). Overall utilization of OR plan is given by the sum product of the weights of block time w_g^u of the specialty Eqn. (3.2) and individual utilization u_g as shown in Eqn. (3.3).

$$u_g = \frac{\sum_{i=1}^{min(d_g, N_g)} p_i}{B_g} \tag{3.1}$$

$$w_g^u = \frac{B_g}{\sum_{g=1}^{G} B_g} \tag{3.2}$$

$$U = \sum_{g=1}^{G} w_g^u \, u_g \tag{3.3}$$

Service level of a specialty s_g is defined as the ratio of the number of cases served to the planned number of cases. Service level of specialty is calculated using Eqn. (3.4). Overall service level of OR plan is obtained from the sum product of weightages of surgical-mix w_g^s of the specialty Eqn. (3.5) and individual service level s_g as shown in Eqn. (3.6).

$$s_g = \frac{min(d_g, N_g)}{N_g} \tag{3.4}$$

$$w_g^s = \frac{N_g}{\sum_{g=1}^{G} N_g} \tag{3.5}$$

$$S = \sum_{g=1}^{G} w_g^s \, s_g \tag{3.6}$$

3.2 Trade-off between utilization and service level

The decision regarding the OR configuration can be made from two different perspectives: (1) Workload load perspective, (2) Demand perspective. These approaches provide optimal OR configuration to maximize utilization and service level respectively.

3.2.1 Maximizing Utilization

To maximize utilization of OR, the block time should be matched to the expected workload of a specialty: the product of demand and processing times. Workload of specialty can be estimated using the joint distribution of historical distribution of demands $d_g \sim \mathcal{N}(\mu_g^d, {\sigma_g^d}^2)$ and processing times $p_g \sim \mathcal{N}(\mu_g^p, {\sigma_g^p}^2)$. The mean μ_g^l and variance $(\sigma_g^l)^2$ of work load are given by Eqn. (3.7) and Eqn. (3.8) respectively.

$$\mu_g^l = \mu_g^d \mu_g^p \tag{3.7}$$

$$(\sigma_g^l)^2 = {\mu_g^d}^2 (\sigma_g^p)^2 + {\mu_g^p}^2 (\sigma_g^d)^2 + (\sigma_g^p)^2 (\sigma_g^d)^2 \tag{3.8}$$

Then, the newsvendor model is employed to obtain optimal block times with the cost ratio of idletime and overtime $\left(\frac{C_i}{C_o}\right)$. The optimal Block time to maximize utilization is given by Eqn. (3.9), similarly, Optimal processing time p_g^* for each case is given by Eqn. (3.10).

$$B_g^u = \mu_g^l + z\sigma_g^l \tag{3.9}$$

Where, $F\left(B_g^u\right) = \Phi(\mathrm{z}) = \frac{C_o}{C_o+C_i}$, and $z = \frac{B_g^u-\mu_g^l}{\sigma_g^l}$.

$$p_g^* = \mu_g^p + z\sigma_g^p \tag{3.10}$$

Surgical-mix N_g^u to maximize utilization is obtained by dividing the allocated block time for the specialty by the optimal processing time p_g^* Eqn. (3.11).

$$N_g^u = \frac{B_g^u}{p_g^*} \tag{3.11}$$

3.2.2 Maximizing Service level

To maximize service level of OR, the surgical-mix should be matched to the expected demand for a specialty. Alike in section 3.2.1, Optimum surgical-mix N_g^s , using newsvendor model is estimated by Eqn. (3.12). Optimal block time for a specialty to maximizes service level B_g^s is obtained by the product of surgical-mix from Eqn. (3.12) and optimal processing time from Eqn. (3.10).

$$N_g^s = \mu_g^d + \mathrm{z}\sigma_g^d \tag{3.12}$$

$$B_g^s = N_g^s p_g^* \tag{3.13}$$

Though the objectives seem to be consistent with each other; there exists a trade-off between them. When the OR configurations from these two perspectives are evaluated in terms of utilization and service level, we can observe that the optimal solution for maximizing utilization is not optimal for maximizing service level.

Table 3. 1: Trade-off between utilization and service level

Criteria	Utilization	Service level
Workload (Maximizes utilization)	U^{u*}	S^{u}
Demand (maximizes service level)	U^{s}	S^{s*}

The elements in Table.3.1 are presented in the form of $KPI^{obj^{opt\,(*)}}$. It can be inferred that, utilization in maximum when the OR configuration is derived from workload a perspective, but the service level is not optimal. Service level is maximum when the OR configuration is derived from demand perspective, but the utilization is not optimal. Therefore, further optimization is needed to minimize the trade-off between utilization and service level.

3.3 Minimizing the Trade-off

Optimal Block times and surgical-mixes are generated explicitly for each specialty in the previous section (section.3.2). We minimize the trade-off between utilization and service level using historical data of OR performance with multiple objective portfolio optimization.

An analogy can be drawn between choosing different stocks with weights in a portfolio, to selecting a surgical-mix and block times of surgical specialties for OR planning. Efficient frontiers of utilization and service level can be formulated using the historical sample mean and standard deviations of utilizations and service levels of surgical specialties.

3.3.1 Efficient frontier of utilization

Utilizations u_g are evaluated using Eqn. (3.1) for each individual specialty over a long period of time for an implemented OR configuration (B_g, N_g). Alike portfolio optimization in Markovitz (1952) we define historical sample mean of utilization $\overline{\mu}_g^u$ as return value of a specialty g and σ_{ij}^u is the covariance of i^{th} and j^{th} specialty. Efficient frontier of utilization are derived from Eqn. (3.14-3.18) which gives the percentage of total block time for each specialty x_g, their respective Expected utilization E_u and the associated risk σ_u : standard deviation of expected utilization.

$$\text{Maximize: } E_u = \sum_{g=1}^{G} x_g \overline{\mu}_g^u \tag{3.14}$$

Subject to:

$$\sum_{g=1}^{G} x_g = 1 \tag{3.15}$$

$$\left(\frac{min(u_g)}{\sum_{g=1}^{G} \overline{\mu_g^l}}\right) \leq x_g \leq \left(\frac{max(u_g)}{\sum_{g=1}^{G} \overline{\mu_g^l}}\right) \tag{3.16}$$

$$\sqrt{\sum_{i=1}^{G} \sum_{j=1}^{G} \sigma_{ij}^u x_i x_j} \leq \sigma_u \tag{3.17}$$

$$x_g \geq 0 \tag{3.18}$$

In the above mathematical formulation, Eqn. (3.14) maximizes the expected utilization subject to constraints Eqn. (3.15-3.18). Constraint (3.15) ensures that the sum of all the weights is equal to one. Constraint (3.16) warrants that the weights of a specialty lie within the maximum and minimum limits. Constraint (3.17) ensures the standard deviation of expected utilization is less than or equal to the associated risk. Constraint (3.18) is to make sure all weights are greater than or equal to zero.

3.3.2 Efficient frontier of service level

A similar approach from section 3.3.1 is extended to the service level. Service levels s_g are evaluated using Eqn. (3.4) for each individual specialty over a long period of time for an implemented OR configuration (B_g, N_g). By defining historical sample mean of service level $\overline{\mu}_g^s$ as return value of a specialty g and σ_{ij}^s is the covariance of i^{th} and j^{th} specialty. Efficient frontier of service level is derived from Eqn. (3.19-3.18) which gives the percentage of total block time for each specialty y_g, their respective Expected utilization E_s and the associated risk σ_s : standard deviation of expected utilization.

$$\text{Maximize: } E_s = \sum_{g=1}^{G} y_g \overline{\mu}_g^s \tag{3.19}$$

Subject to:

$$\sum_{g=1}^{G} y_g = 1 \tag{3.20}$$

$$\left(\frac{min(s_g)}{\sum_{g=1}^{G} \overline{\mu_g^l}}\right) \leq x_g \leq \left(\frac{max(s_g)}{\sum_{g=1}^{G} \overline{\mu_g^l}}\right) \tag{3.21}$$

$$\sqrt{\sum_{i=1}^{G}\sum_{j=1}^{G}\sigma_{ij}^{s}y_i y_j} \leq \sigma_s \tag{3.22}$$

$$y_g \geq 0 \tag{3.23}$$

In the above mathematical formulation, Eqn. (3.19) maximizes the expected utilization subject to constraints Eqn. (3.20-3.23). Constraint (3.20) ensures that the sum of all the weights is equal to one. Constraint (3.21) warrants that the weights of a specialty lie within the maximum and minimum limits. Constraint (3.22) ensures the standard deviation of expected service level is less than or equal to the associated risk. Constraint (3.23) is to make sure all weights are greater than or equal to zero.

3.3.3 Efficient frontier to minimize the trade-off

To minimize the trade-off between utilization and service level. We used multiple objective portfolio optimization method which is similar weighted sum multiple objective optimization. We normalize the multi-objective function, by dividing it with the ranges of expected objective function values derived with single objectives from section 3.3.1 and 3.3.2. Ranges of Expected utilization and service level are the difference between the upper and lower bounds of E_u and E_s respectively. Similarly, ranges of associated risks are the difference between the upper and lower bounds of σ_u and σ_s respectively. The objective of the multiple -objective portfolio optimization is to minimize the deviation from the optimal solutions of objectives, given the constraints and preference among the objectives.

Efficient frontier to minimize the trade-off from the mathematical formulation Eqn. (3.24-3.30) which gives the percentage of total block time for each specialty x_g^{α}, percentage of surgical mix for each specialty y_g^α with their respective Expected trade-off value E^α and the associated risk in trade-off σ^α: standard deviation of expected trade-off value.

$$\text{Minimize: } E^\alpha = \frac{\alpha}{\overline{E}_U - \underline{E}_U}\left(\overline{E}_U - \sum_{g=1}^{G} x_g^\alpha \overline{\mu_g^u}\right) + \frac{(1-\alpha)}{\overline{E}_s - \underline{E}_s}\left(\overline{E}_s - \sum_{g=1}^{G} y_g^\alpha \overline{\mu_g^s}\right) \tag{3.24}$$

Subject to:

$$\sum_{g=1}^{G} x_g^\alpha = 1 \tag{3.25}$$

$$\sum_{g=1}^{G} y_g^\alpha = 1 \tag{3.26}$$

$$\left(\frac{min(u_g)}{\sum_{g=1}^{G} \overline{\mu_g^u}}\right) \leq x_g^\alpha \leq \left(\frac{max(u_g)}{\sum_{g=1}^{G} \overline{\mu_g^u}}\right) \tag{3.27}$$

$$\left(\frac{min(s_g)}{\sum_{g=1}^{G} \overline{\mu_g^s}}\right) \leq y_g^\alpha \leq \left(\frac{max(s_g)}{\sum_{g=1}^{G} \overline{\mu_g^s}}\right) \tag{3.28}$$

$$y_g^\alpha \left(\sum_{g=1}^{G} N_g^d\right) p_g \leq x_g^\alpha \left(\sum_{g=1}^{G} B_g^l\right) \ \forall\ g = 1,2..G \tag{3.29}$$

$$\frac{\alpha}{\overline{\sigma}_u - \underline{\sigma}_u}\left(\overline{\sigma}_u - \sqrt{\sum_{i=1}^{G}\sum_{j=1}^{G} \sigma_{ij}^u x_i x_j}\right) + \frac{(1-\alpha)}{\overline{\sigma}_s - \underline{\sigma}_s}\left(\overline{\sigma}_s - \sqrt{\sum_{i=1}^{G}\sum_{j=1}^{G} \sigma_{ij}^s y_i y_j}\right) \geq \sigma^\alpha \tag{3.30}$$

In the above mathematical formulation, Eqn. (3.19) Minimizes the trade-off among the objectives, when preference among the objectives α is given. The objective function is Subjected to constraints Eqn. (3.25-3.30). Constraint (3.25) and (3.26) ensures that the sum of all the weights is equal to one for the block times and surgical mix respectively. Constraint (3.27) warrants that the weights of a specialty lie within the maximum and minimum workloads. Constraint (3.28) warrants that the weights of a specialty lie within the maximum and minimum limits of demand. Constraint (3.29) is to ensure that the product of surgical mix and processing time is less than or equal to allocated block time.

Constraint (3.30) ensures the standard deviation of expected trade-off value is less than or equal to the associated risk.

The expected utilization E_u^α and associated risk σ_u^α for a preference can be evaluated by Eqn. (3.31) and Eqn. (3.32) respectively. The expected service level E_s^α and associated risk σ_s^α for a preference can be evaluated by Eqn. (3.33) and Eqn. (3.34) respectively.

$$E_u^\alpha = \sum_{g=1}^{G} x_g^\alpha \overline{\mu_g^u} \tag{3.31}$$

$$\sigma_u^\alpha = \sqrt{\sum_{i=1}^{G}\sum_{j=1}^{G} \sigma_{ij}^u x_i^\alpha x_j^\alpha} \tag{3.32}$$

$$E_s^\alpha = \sum_{g=1}^{G} y_g^\alpha \overline{\mu_g^s} \tag{3.33}$$

$$\sigma_s^\alpha = \sqrt{\sum_{i=1}^{G}\sum_{j=1}^{G} \sigma_{ij}^s y_i^\alpha y_j^\alpha} \tag{3.34}$$

OR configurations (B_g^α, N_g^α) which minimize trade-off between utilization and service rate is obtained by multiplying the weights of utilization x_g^α with the sum of block times from Eqn. (3.9) and the surgical-mix is obtained by multiplying the weights of service level y_g^α with the sum of cases from Eqn. (3.12) respectively for preference α.

$$B_g^{\alpha} = \left(\sum_{g=1}^{G} B_g^{u}\right) x_g^{\alpha} \tag{3.35}$$

$$N_g^{\alpha} = \left(\sum_{g=1}^{G} N_g^{d}\right) y_g^{\alpha} \tag{3.36}$$

To sum up, analyzing the distribution of demand and processing times of each surgical specialty, OR configurations are modeled minimizing the trade-off between utilization and service. A one to one relationship is provided relating the distributions of demand, processing times, OR configurations and preference among objectives, to the distribution of expected utilization and service level as shown in the Figure 3.1.

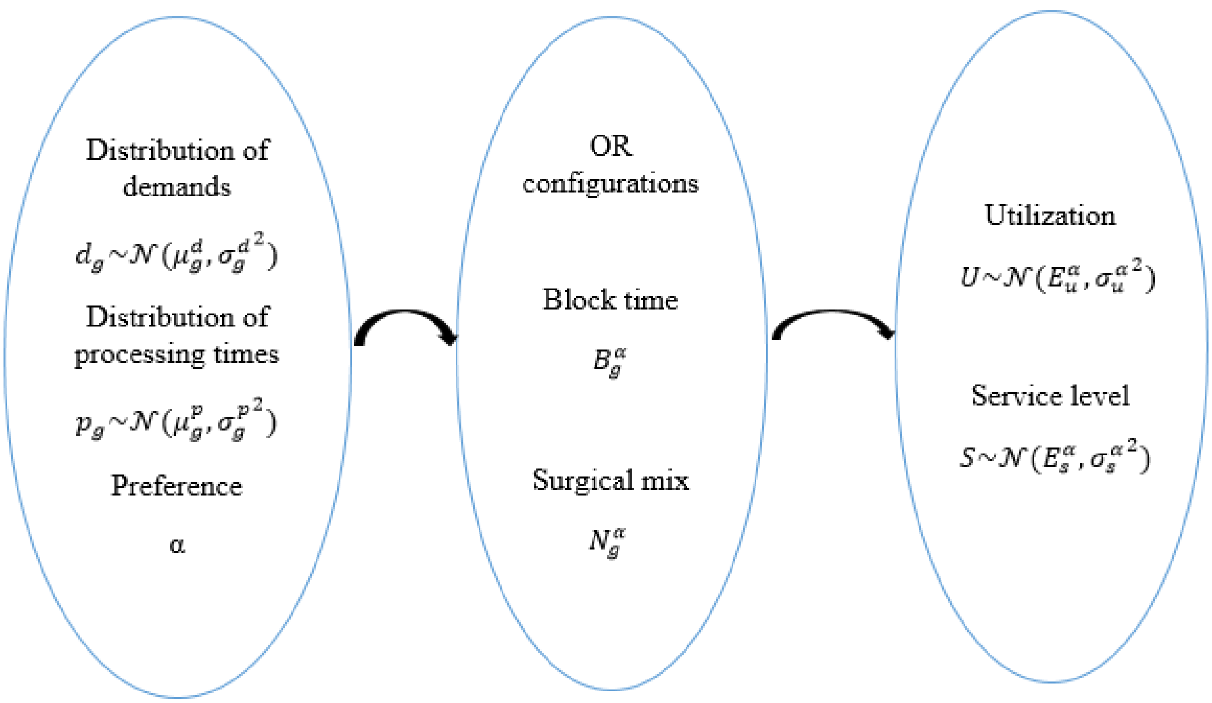

Figure 3. 1:One to one relationship

This model of one to one mapping aides OR managers to make an informed decision about the OR configuration according to their targets of OR performances and preferences.

3.4 Adaptive control scheme

OR managers are required to maintain the OR performance within controllable limits to meet their targets in the long term. This enhances the need to develop an adaptive control scheme, which tracks the OR performance along the time horizon and changes the OR configuration periodically to meet the targets in the long term. Uniform OR configurations are repeatedly implemented for a limited period of time: referred as cycle c. If the OR performance at the end of each cycle is within the predetermined control limits then the OR performance, then the system is under control. On the other hand, if the OR performance is out of the control, then the OR configuration must be modified using an adaptive control scheme to bring the OR performance in control.

As previous section (3.3.3) provides one to one relationship between the OR configurations and OR performance distributions, it can be used by the OR managers to decide on OR configurations depending on their, preference among the objectives, targets of utilization and service level. A decision on OR configuration B_g^α, N_g^α results in utilization and service level following the distribution $U \sim \mathcal{N}(E_u^\alpha, {\sigma_u^\alpha}^2)$ and $S \sim \mathcal{N}(E_s^\alpha, {\sigma_s^\alpha}^2)$ respectively. For a cycle c, upper and lower control limits for OR utilization are given by Eqn. (3.37) and Eqn. (3.38) respectively, where β is the fraction of standard deviation allowed. Similarly, Controllable upper limit and lower limits for OR service levels are Eqn. (3.39) and Eqn. (3.40) respectively.

$$\overline{E}_U^c = E_u^\alpha + \beta\sigma_u^\alpha \tag{3.37}$$

$$\underline{E}_U^c = E_u^\alpha - \beta\sigma_u^\alpha \tag{3.38}$$

$$\overline{E}_S^c = E_s^\alpha + \beta\sigma_s^\alpha \tag{3.39}$$

$$\underline{E}_S^c = E_S^{\alpha} - \beta\sigma_S^{\alpha} \tag{3.40}$$

Actual OR utilization X_U^c and service level X_S^c are evaluated at the end of each cycle using Eqn. (3.3) and Eqn. (3.6) respectively. As shown in Fig.4, If utilization and service level are within the controllable limits, then same OR configuration is continued for the following cycle. Else, OR configuration for the next cycle is derived by changing the expected utilization, service level and preference Eqn. (3.41) among the objectives for the next cycle.

If $\underline{E}_U^c \le X_U^c \le \overline{E}_U^c$ **and** $\underline{E}_S^c \le X_S^c \le \overline{E}_S^c$

$$E_U^{c+1} = E_U^c$$

$$E_S^{c+1} = E_S^c$$

Else

$$E_U^{c+1} = 2E_U^c - X_U^c$$

$$E_S^{c+1} = 2E_S^c - X_S^c$$

End

Figure 3. 2:Adaptive scheme

$$\alpha^{c+1} = \frac{\left(\frac{X_S^c}{E_S^C}\right)}{\left(\frac{X_U^C}{E_U^C} + \frac{X_S^c}{E_S^C}\right)} \tag{3.41}$$

OR configurations for the next cycle are formulated using the weights of block times x_g^c and surgical mix y_g^c with preference, expected utilization and service level for next cycle as a constraint in the mathematical formulation mentioned in Eqn. (3.42-49)

$$\text{Maximize: } \sigma^{c+1} = \left[\frac{\alpha^{c+1}}{\overline{\sigma}_u - \underline{\sigma}_u}\left(\overline{\sigma}_u - \sqrt{\sum_{i=1}^{G}\sum_{j=1}^{G}\sigma_{ij}^{u}x_i^{c+1}x_j^{c+1}}\right) + \frac{(1-\alpha^{c+1})}{\overline{\sigma}_s - \underline{\sigma}_s}\left(\overline{\sigma}_s - \sqrt{\sum_{i=1}^{G}\sum_{j=1}^{G}\sigma_{ij}^{s}y_i^{c+1}y_j^{c+1}}\right)\right] \quad (3.42)$$

Subject to:

$$\sum_{g=1}^{G} x_g^{c+1} = 1 \quad (3.43)$$

$$\sum_{g=1}^{G} y_g^{c+1} = 1 \quad (3.44)$$

$$\left(\frac{min(u_g)}{\sum_{g=1}^{G}\overline{\mu_g^u}}\right) \leq \ x_g^{c+1} \leq \left(\frac{max(u_g)}{\sum_{g=1}^{G}\overline{\mu_g^u}}\right) \ \forall \ g = 1,2..G \quad (3.45)$$

$$\left(\frac{min(s_g)}{\sum_{g=1}^{G}\overline{\mu_g^s}}\right) \leq \ y_g^{c+1} \leq \left(\frac{max(s_g)}{\sum_{g=1}^{G}\overline{\mu_g^s}}\right) \ \forall \ g = 1,2..G \quad (3.46)$$

$$y_g^{c+1}\left(\sum_{g=1}^{G} N_g^d\right)p_g \leq x_g^{c+1}\left(\sum_{g=1}^{G} B_g^l\right) \ \forall \ g = 1,2..G \quad (3.47)$$

$$\sum_{g=1}^{G} x_g^{c+1}\overline{\mu_g^u} \leq E_U^{c+1} \quad (3.48)$$

$$\sum_{g=1}^{G} y_g^{c+1}\overline{\mu_g^s} \leq E_S^{c+1} \quad (3.49)$$

In the above mathematical formulation, Eqn. (3.42) Maximizes the deviation from the maximum risk of utilization and service levels, when preference among the objectives α^{c+1} is given. The objective function is Subjected to constraints Eqn. (3.43-49). Constraint (3.43) and (3.44) ensures that the sum of all the weights is equal to one for the block times and surgical mix respectively. Constraint (3.45) warrants that the weights of a specialty lie within the maximum and minimum workloads. Constraint (3.46) warrants that the weights of a specialty lie within the maximum and minimum limits of demand. Constraint (3.47) is to ensure that the product of surgical mix and processing time is less than or equal to allocated block time. Constraint (3.48) and (3.49) ensures the expected utilization and service levels are less than prescribed limits. The block time and surgical mix for the next cycle are given with the weights x_g^{c+1} and y_g^{c+1} substituted in Eqn. (3.35) and Eqn. (3.36) respectively.

Chapter 4 Case study

This chapter provides the results of the case study. A brief background of the UKHC is introduced. First, utilization and service level are compared among two sets of OR configuration each developed from demand and workload perspective, to show the trade-off among these KPI's of OR. Second, efficient portfolio frontiers to minimize the trade-off between utilization and service level are generated and compared. Third, the adaptive control scheme is validated by verifying the conformance of utilization and service levels within the controllable limits along the time horizon.

University of Kentucky Health Care (UKHC) served almost 30,000 patients from 2013-14, excluding weekends and holidays which is approximately 500 cases in a week. There are 19 major surgical specialties offered at UKHC. Utilization and service level are significant performance indicators at UKHC. Therefore, we intend to study the performance of OR regarding utilization and service level along the time horizon. From the historical data, we have the number of cases of each surgical group served and processing times, of each week over a period of one year. Data analysis showed that these groups followed a normal distribution regarding both number of cases processed per week and processing times, with a confidence interval higher than 95%.

Table 4. 1: Distribution of demand and processing time at UKHC

	Demands		**Case times**	
	Mean	Std.Dev	Mean	Std.Dev
service-1	28.15	7.95	141.02	19.01
service-2	30.64	7.82	248.16	28.64
service-3	6.91	2.36	129.02	18.84
service-4	54.51	10.66	116.56	18.40
service-5	24.25	8.22	171.17	23.95
service-6	25.26	8.07	135.38	20.41
service-7	10.68	3.84	123.61	27.00
service-8	15.68	4.12	171.93	25.99
service-9	31.94	8.26	187.77	20.40
service-10	8.04	2.71	170.33	40.64
service-11	49.34	14.37	64.22	12.67
service-12	95.53	22.25	143.01	11.20
service-13	20.96	6.27	88.44	16.39
service-14	38.85	9.89	141.19	19.82
service-15	4.02	2.39	80.93	25.17
service-16	24.08	9.90	103.31	16.40
service-17	4.62	2.75	226.78	75.78
service-18	34.17	9.23	120.47	17.00
service-19	15.34	4.45	166.74	31.68

4.1 Trade-off between utilization and service level

Simulation is carried out with randomly generated normal demands and processing times at discrete time intervals. The performance of OR configurations is measured in terms of utilization and service level. Simulation results of the OR configurations derived using newsvendor model from workload and demand perspectives are presented in Table:4.2.

Table 4. 2: Case study - Trade-off between utilization and service rate

	utilization	Service level
Workload perspective	**0.831***	0.843

Demand perspective	0.814	**0.865***

Table 4.2 clearly presents the trade-off among the OR configurations from workload and demand perspectives. Configuration from workload perspective has high utilization. On the other hand, configurations from demand perspective have high service rate.

4.2 Minimizing the trade-off

A configuration which is optimal on both utilization and service rate is obtained by minimizing the tradeoff using PS. Optimal configuration minimizing the trade-off should have minimum possible expected value of trade-off and maximum possible risk. From Figure 4.1, we observe that, full preference ($\alpha = 1$) to utilization offers minimum trade-off, but also minimum risk. On the other hand, with full preference to service level ($\alpha = 0$) offers maximum expected trade-off associated with maximum risk. A balance between the expected trade-off and risk is achieved with equal preference among objectives ($\alpha = 0.5$).

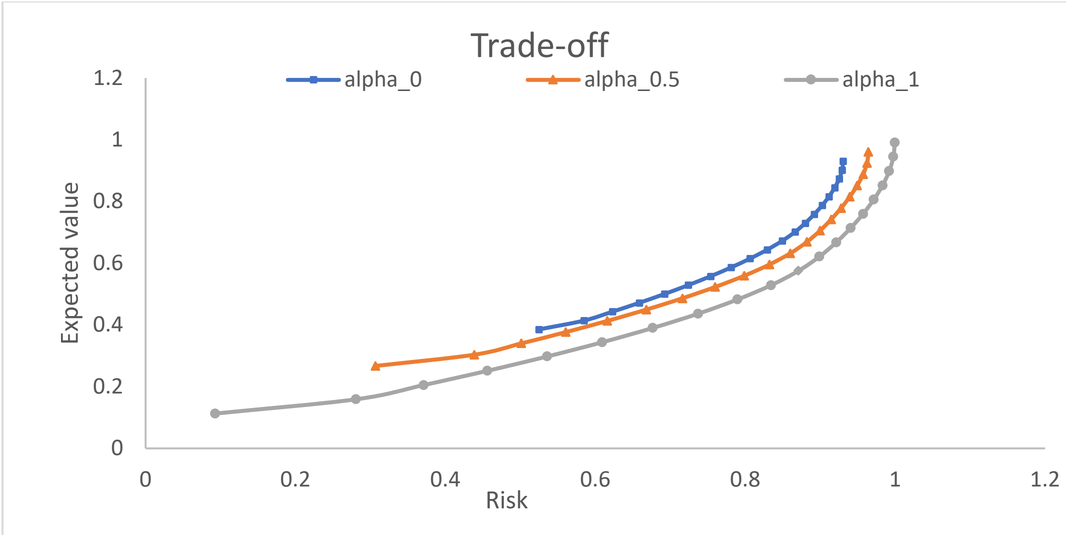

Figure 4. 1:Efficient frontiers of Trade-off

OR configurations minimizing the trade-off are compared in terms of expected utilization and associated standard deviation. From Figure 4.2 it can be observed that efficient frontier generated with full preference to utilization ($\alpha = 1$) is dominating other frontiers.

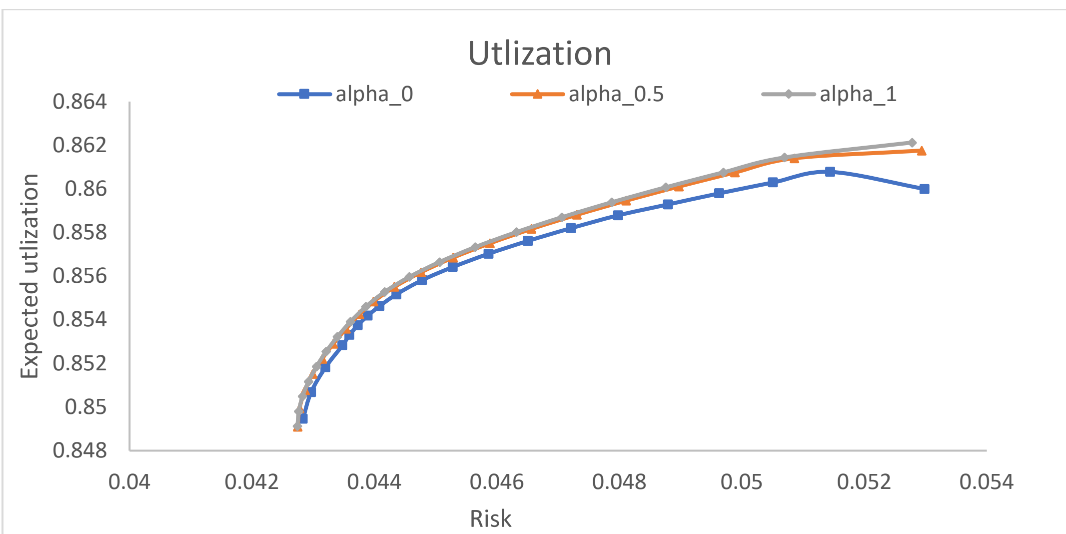

Figure 4. 2: Utilization on efficient frontiers minimizing the trade-off

Similarly, OR configurations minimizing the trade-off are compared in terms of expected service level and associated standard deviation. From Figure 4.3 it can be observed that efficient frontier generated with full preference to service level ($\alpha = 0$) is dominating other frontiers.

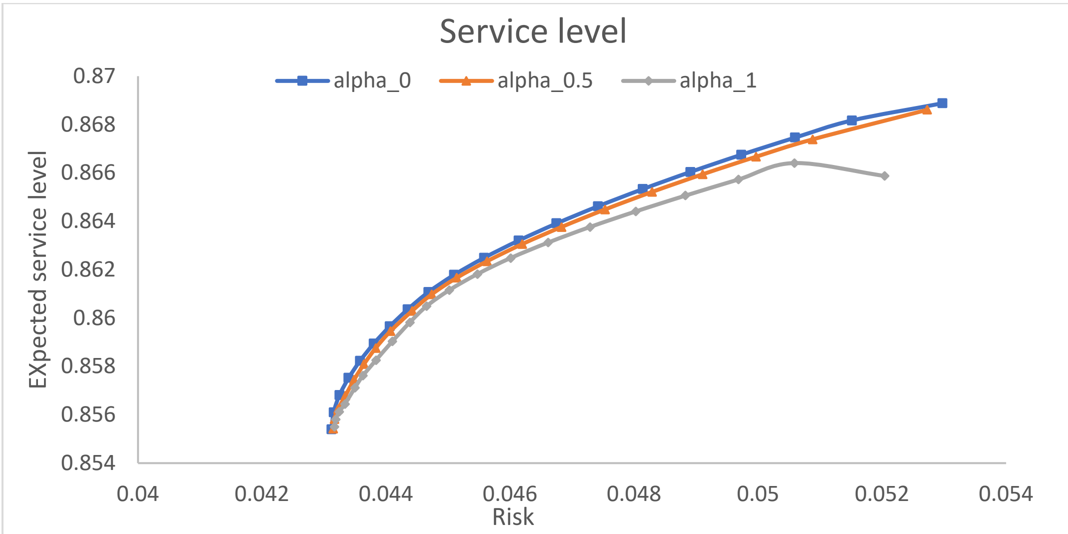

Figure 4. 3: Service level on efficient frontiers minimizing the trade-off

4.3 Adaptive control

Adaptive control model is validated by Simulating normally distributed demands and processing times over period of 50 cycles of 36 weeks each. This model is evaluated at three different preferences among the objectives. Individual control charts of utilization, service level and trade-off values are presented along the time horizon at each cycle. X-bar and R-bar charts, a type of statistical control charts is used to monitor the mean and range of utilization and service level in subgroups of weeks. Capability analysis is carried out to verify the robustness of OR performance measures. A control chart is used to monitor the preference along the time horizon.

4.3.1 Preference: Alpha = 0.5

OR configuration with equal preference among the objectives with minimum expected trade-off value is chosen for the first cycle. OR performance regarding the Utilization and service level are evaluated at the end of each cycle to verify, whether they lie within the controllable limits. Controllable limits for utilization and service levels are established within 0.75 of standard deviation.

Individual X-bar chart of utilization is presented in Figure 4. 4. It can be observed that the utilization is clearly under control. X-bar and R-Chart is used with the continuous data collected in subgroup size of thirty-six. The Mean (X-Bar) of each subgroup for utilization charted on the top graph and the Range (R) of the subgroup of utilization charted on the bottom graph in Figure 4. 5.

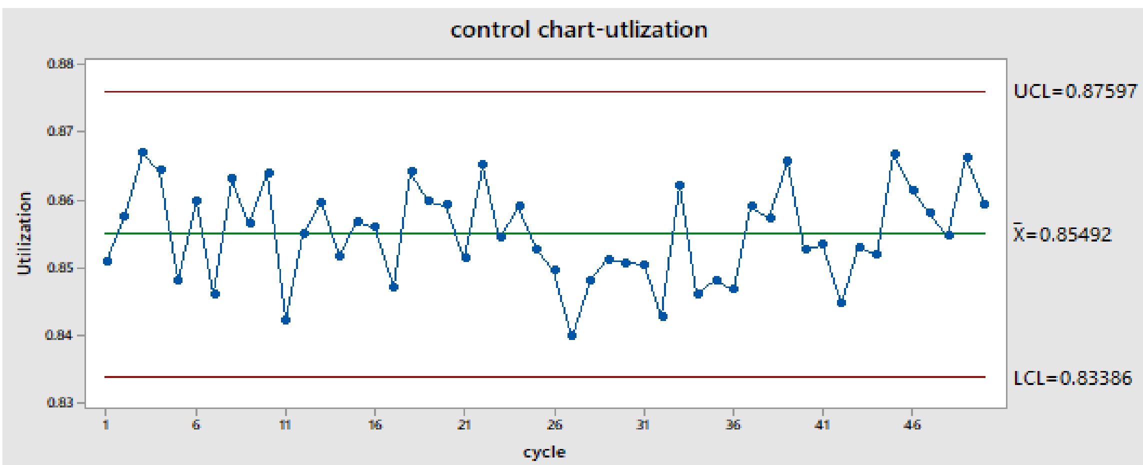

Figure 4. 4:Control chart of utilization with alpha = 0. 5

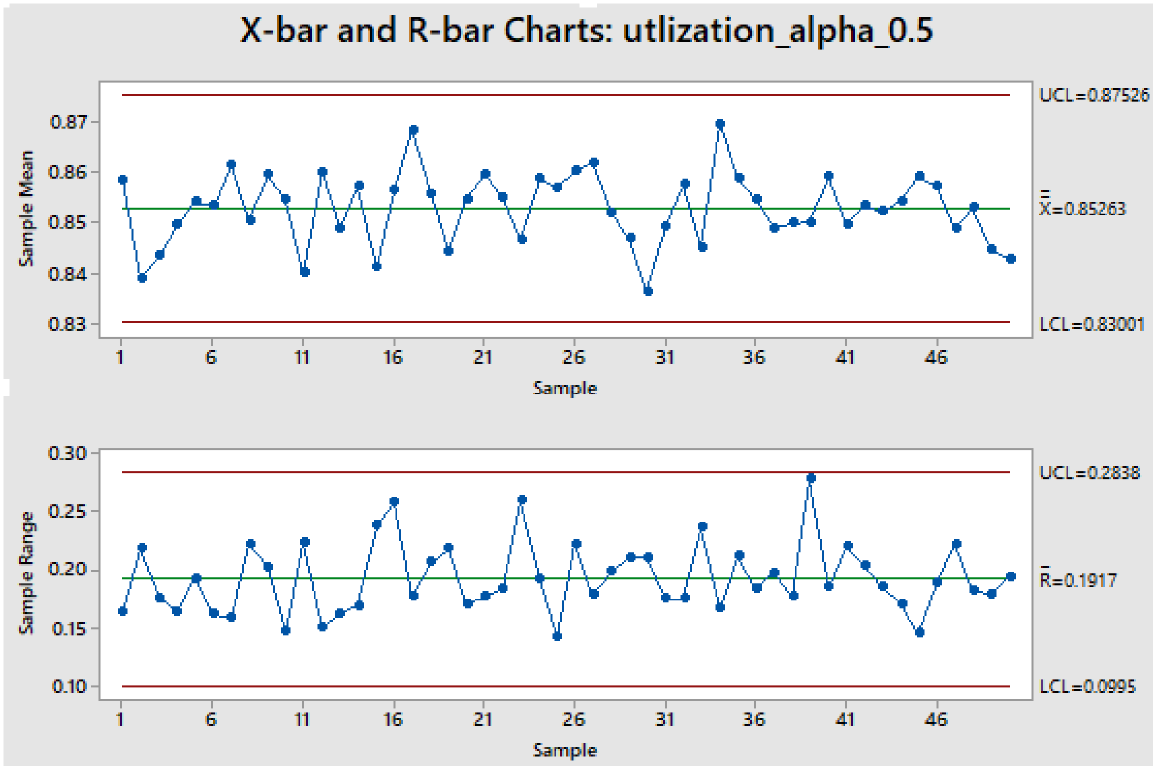

Figure 4. 5:X-bar and R-bar Charts of utilization with alpha = 0. 5

Individual X-bar chart of service level is presented in Figure 4. 6. It can be observed that the service level is clearly under control. X-bar and R-Chart is used with the continuous data collected in subgroup size of thirty-six. The Mean (X-Bar) of each subgroup for service level is charted on the top graph and the Range (R) of the subgroup of service level is charted on the bottom graph in Figure 4. 7.

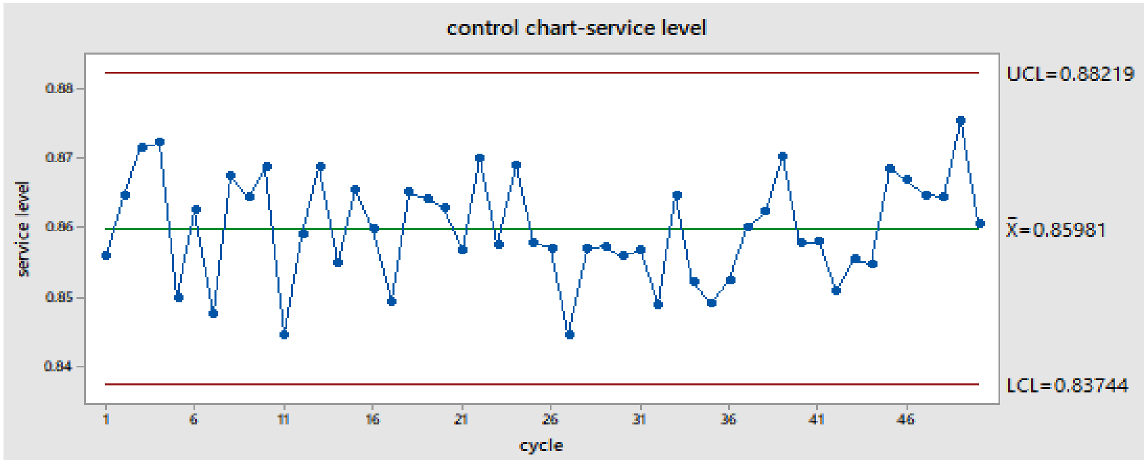

Figure 4. 6: Control chart of service level with alpha = 0. 5

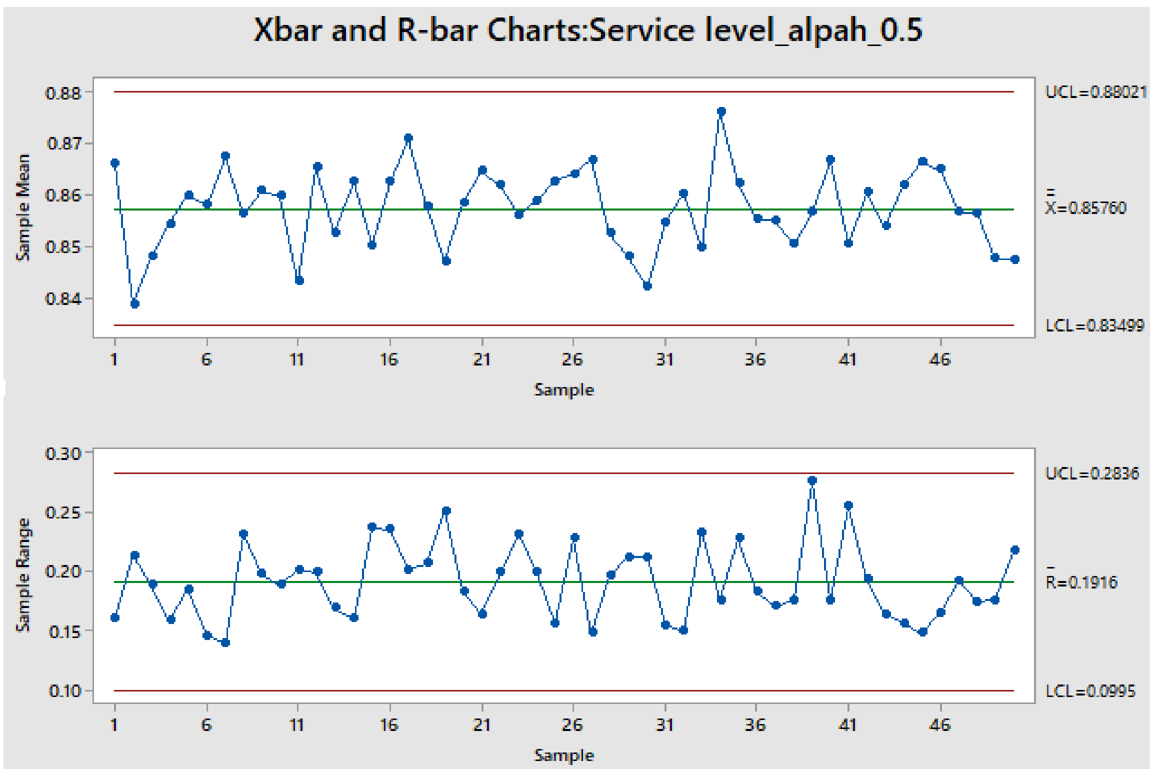

Figure 4. 7:X-bar and R-bar Charts of Service level with alpha = 0. 5

Statistical process control chart of trade-off is monitored along the time, as shown in Figure 4. 8, and is observed to clearly under the control.

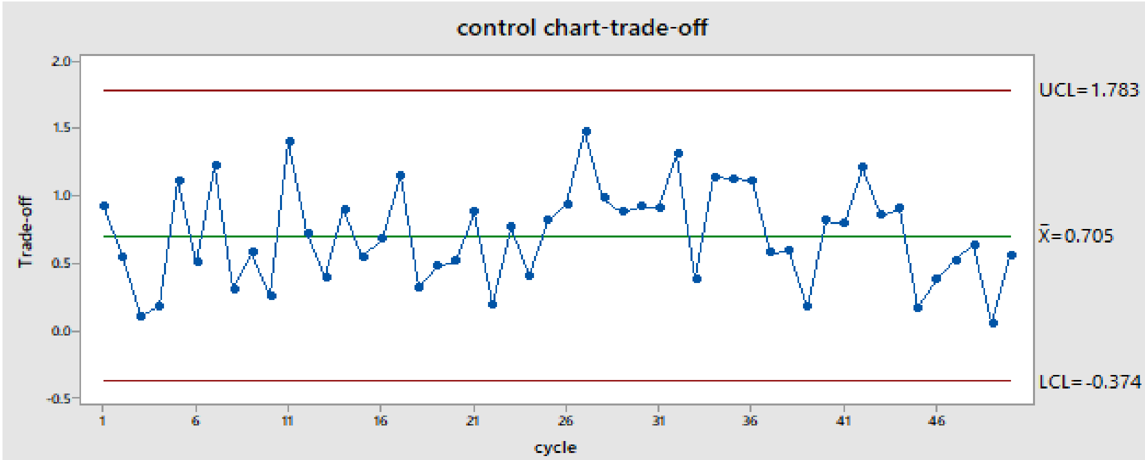

Figure 4. 8:Control chart of Trade-off with alpha = 0. 5

A capability analysis is also carried out to verify, if the control limits of OR performance are within the specified limits. From Figure's 4.9 and 4.10, it is observed that process capability indices $c_p$and c_{pk} values are greater than one, indicating that the OR performance measures: utilization and service levels, are within the pre-established controllable limits.

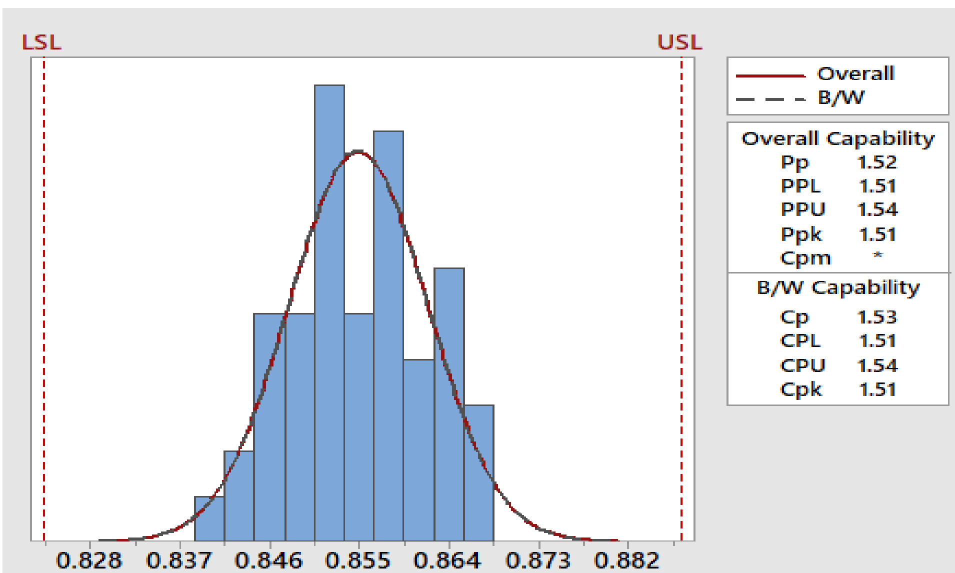

Figure 4. 9:Capability analysis of utilization with alpha = 0. 5

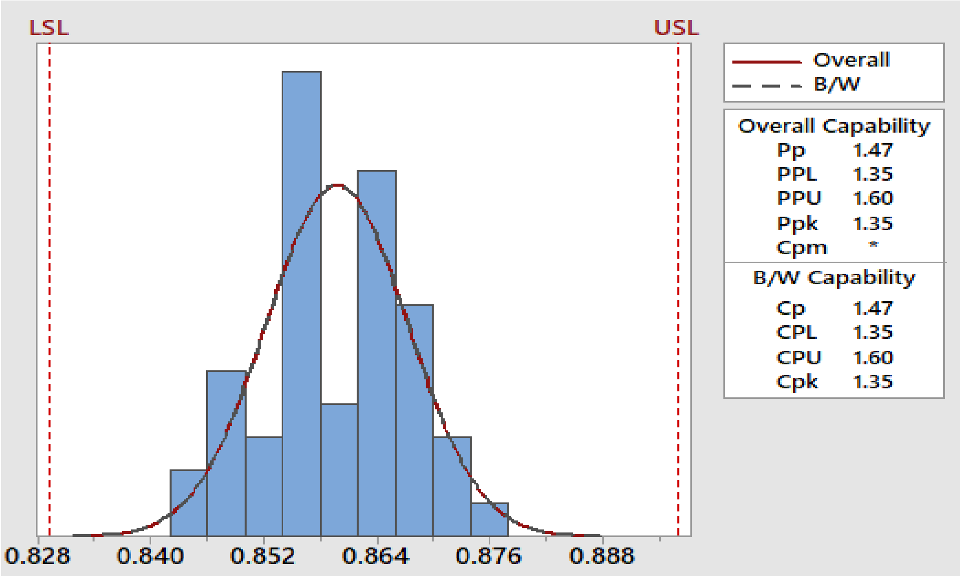

Figure 4. 10:Capability analysis of service level with alpha = 0. 5

Control chart of preference adaptively changing along the cycles is monitored in Figure 4. 11. The mean (X-bar) of the preference is observed to 0.499 approximately equal 0.5, which is the preference chosen initially.

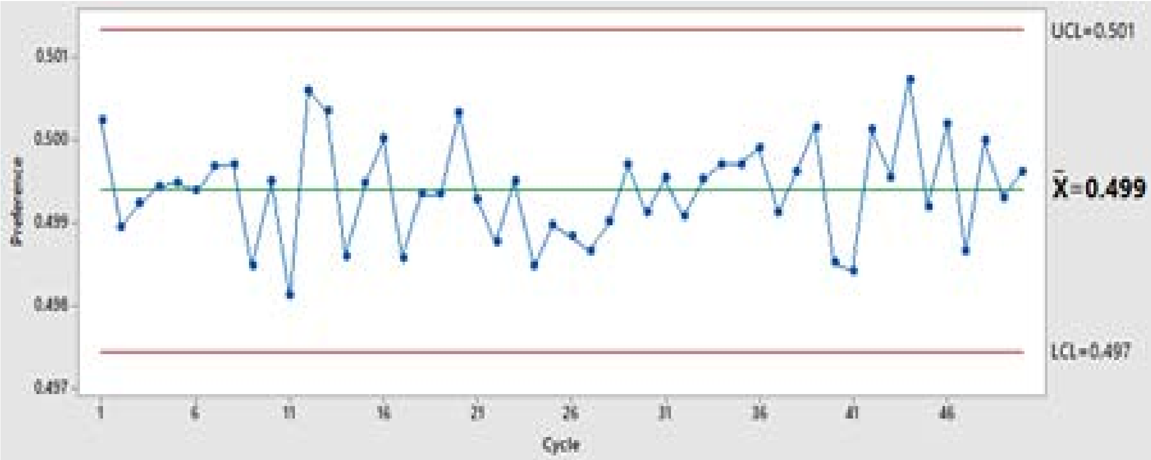

Figure 4. 11:Control chart of alpha with alpha = 0. 5

The promising results from the control charts and process capability analysis show that OR performance regarding utilization and service level are within the control limits along the time horizon. Therefore, it can be assured that our model of optimization and adaptive control enable OR managers to make an informed decision on OR configurations and control the OR performances in the long-term, given the stochastic demand and processing times of surgical specialties.

4.3.2 Preference: Alpha = 0.25

To study the impact of preference on the OR performance measures, we repeated the adaptive control planning with a preference alpha = 0.25 among the objectives: the preference for utilization is 0.25 and the preference of service level is 0.75. Individual X-bar chart of utilization is presented in Figure 4. 12. It can be observed that the utilization is clearly under control and is marginally less than the utilization obtained with equal preference. X-bar and R-Chart is used with the continuous data collected in subgroup size of thirty-six. The Mean (X-Bar) of each subgroup for utilization charted on the top graph and the Range (R) of the subgroup of utilization charted on the bottom graph in Figure 4. 13.

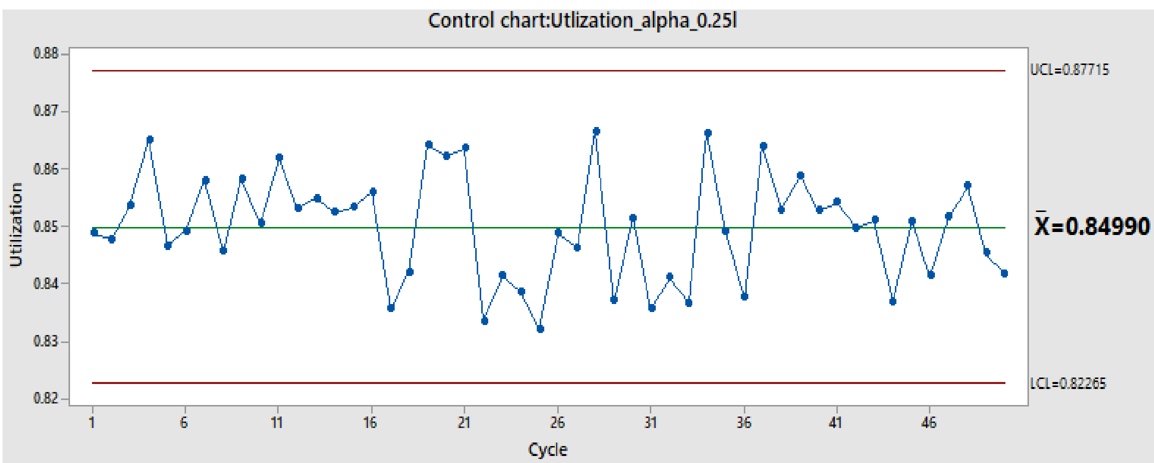

Figure 4. 12:Control chart of utilization with alpha = 0.25

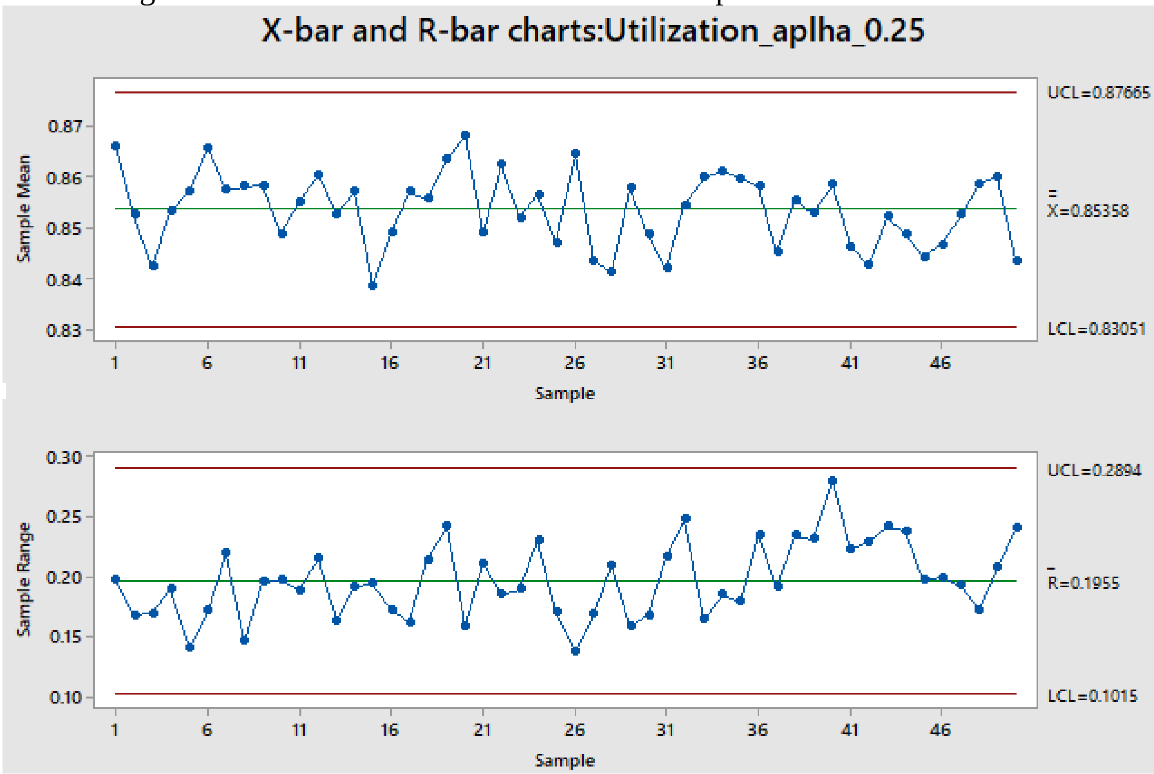

Figure 4. 13:X-bar and R-bar Charts of utilization with alpha = 0.25

Individual X-bar chart of service level is presented in Figure 4. 14. It can be observed that the service level is clearly under control and is marginally greater than the service level with equal preference. X-bar and R-Chart is used with the continuous data collected in subgroup size of thirty-six. The Mean (X-Bar) of each subgroup for service level is charted on the top graph and the Range (R) of the subgroup of service level is charted on the bottom graph in Figure 4. 15.

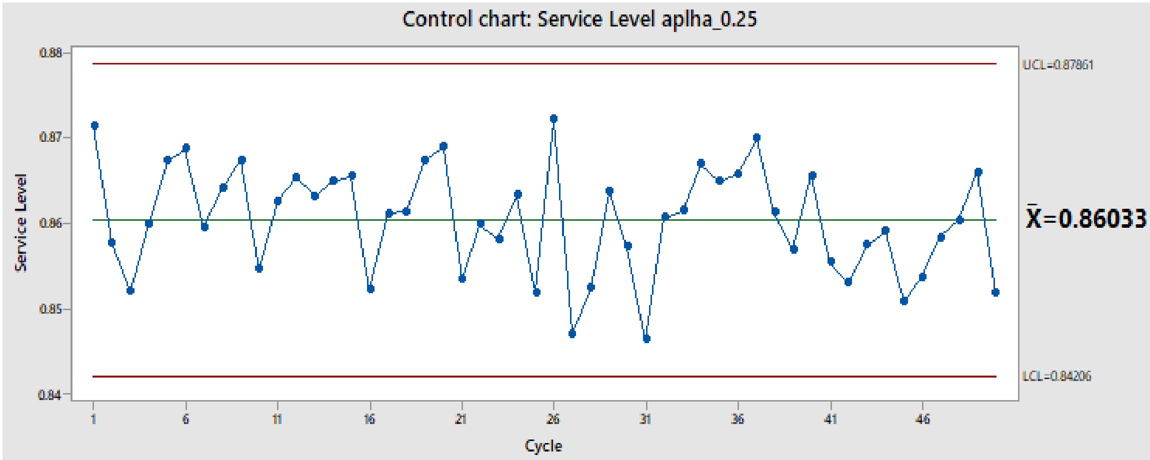

Figure 4. 14:Control chart of Service level with alpha = 0.25

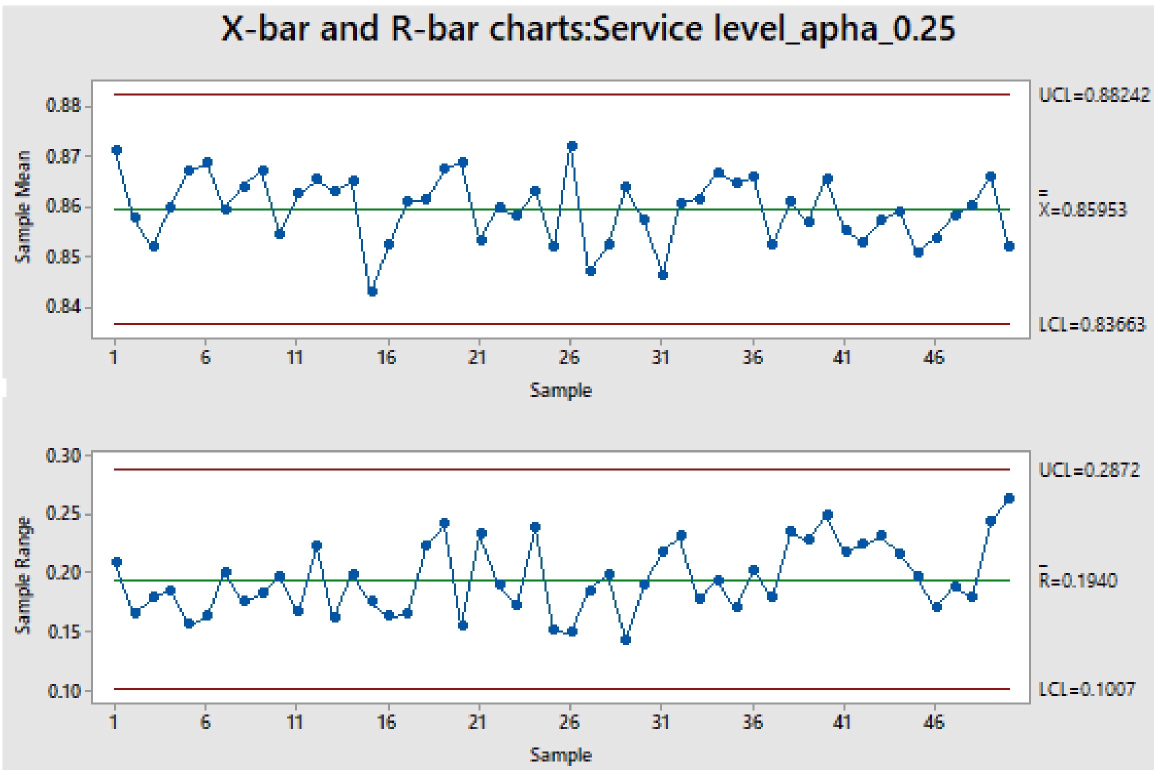

Figure 4. 15:X-bar and R-bar Charts of Service level with alpha = 0.25

Statistical process control chart of trade-off is monitored along the time, as shown in Figure 4. 16, and is observed to clearly under the control with greater trade-off value than equal preference among the objectives.

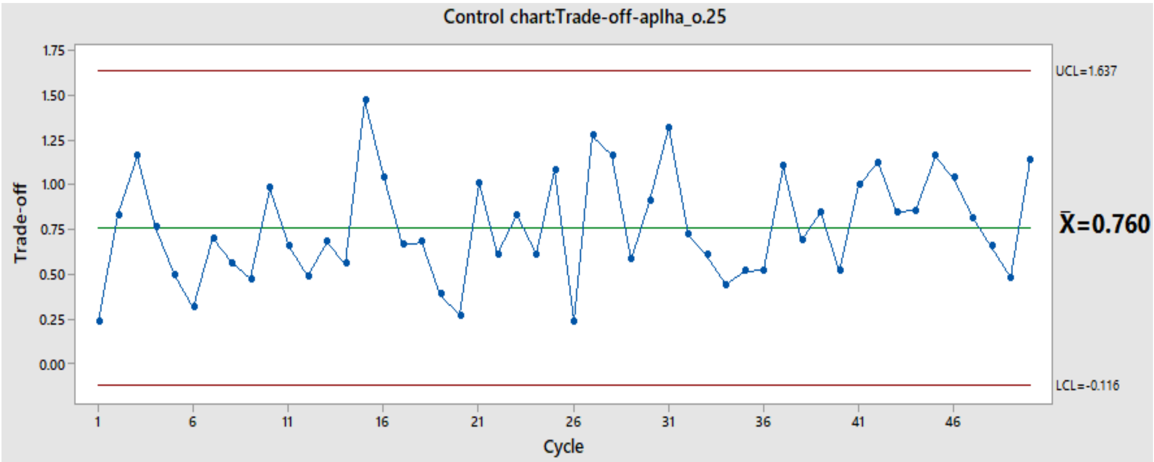

Figure 4. 16:Control chart of Trade-off with alpha = 0.25

Capability analysis is also carried out to verify, if the control limits of OR performance are within the specified limits. From Figure's 4.17 and 4.18, it is observed that process capability indices $c_p$and c_{pk} values are greater than one, indicating that the OR performance measures: utilization and service levels, are within the pre-established controllable limits.

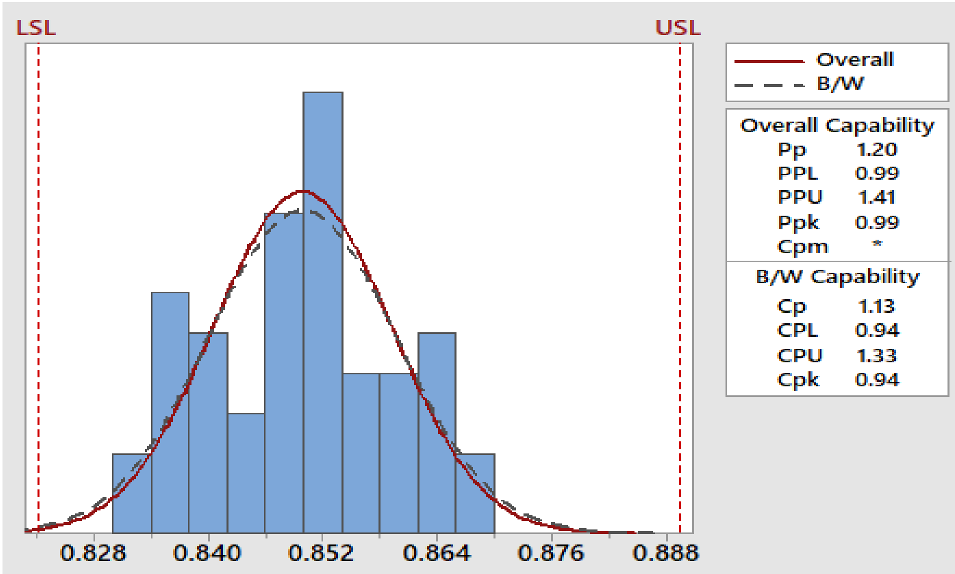

Figure 4. 17:Capability analysis of utilization with alpha = 0.25

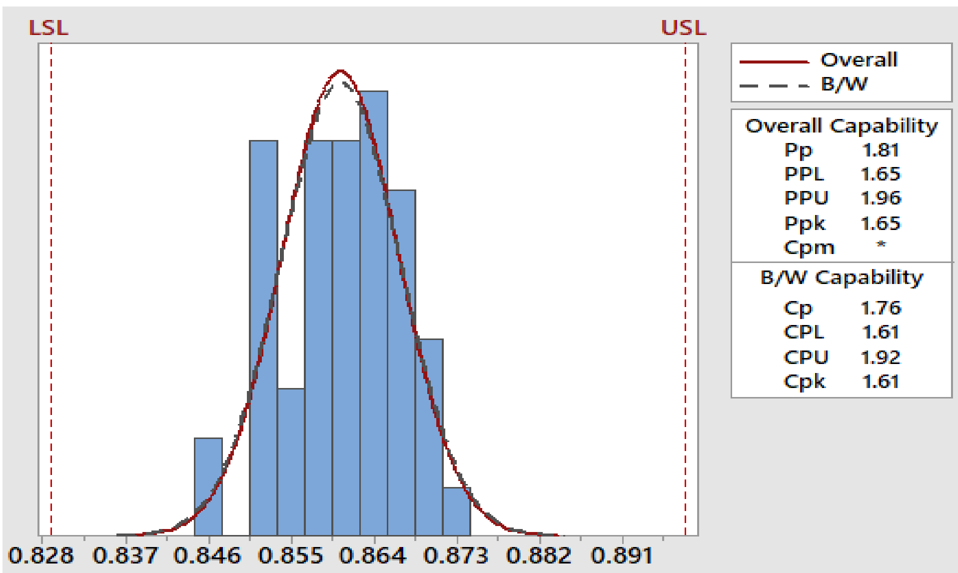

Figure 4. 18:Capability analysis of service level with alpha = 0.25

Control chart of preference adaptively changing along the cycles is monitored in Figure 4. 11. The mean (X-bar) of the preference is observed to 0.499 approximately equal 0.5, which is the preference chosen initially.

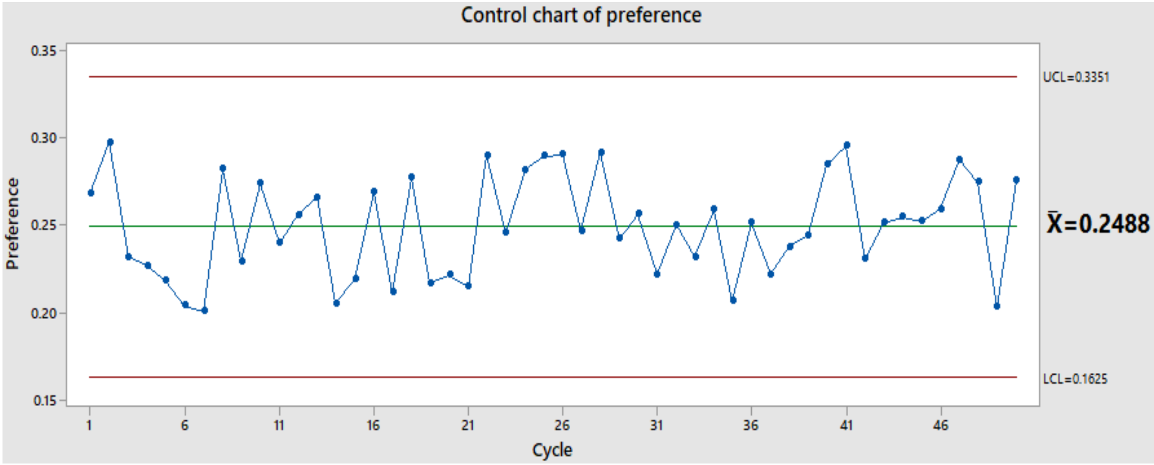

Figure 4. 19:Control chart of preference with alpha=0.25

The promising results from the control charts and process capability analysis show that OR performance regarding utilization and service level are within the control limits along the time horizon. Therefore, it can be assured that our model of optimization and adaptive control enable OR managers to make an informed decision on OR configurations and control the OR performances in the long-term, given the stochastic demand and processing times of surgical specialties.

4.3.3 Preference: Alpha = 0.75

To study the impact of preference on the OR performance measures, we repeated the adaptive control planning with a preference alpha = 0.75 among the objectives: the preference for utilization is 0.75 and the preference of service level is 0.25. Individual X-bar chart of utilization is presented in Figure 4. 20. It can be observed that the utilization is clearly under control and is marginally greater than the utilization obtained with equal preference because of high weightage. X-bar and R-Chart is used with the continuous data collected in subgroup size of thirty-six. The Mean (X-Bar) of each subgroup for utilization charted on the top graph and the Range (R) of the subgroup of utilization charted on the bottom graph in Figure 4. 21.

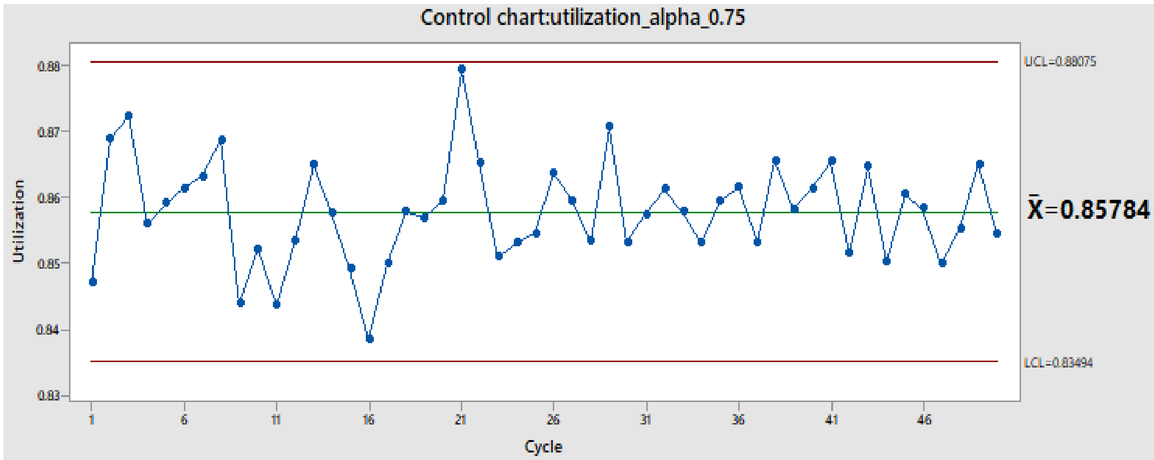

Figure 4. 20:Control chart of utilization with alpha = 0.75

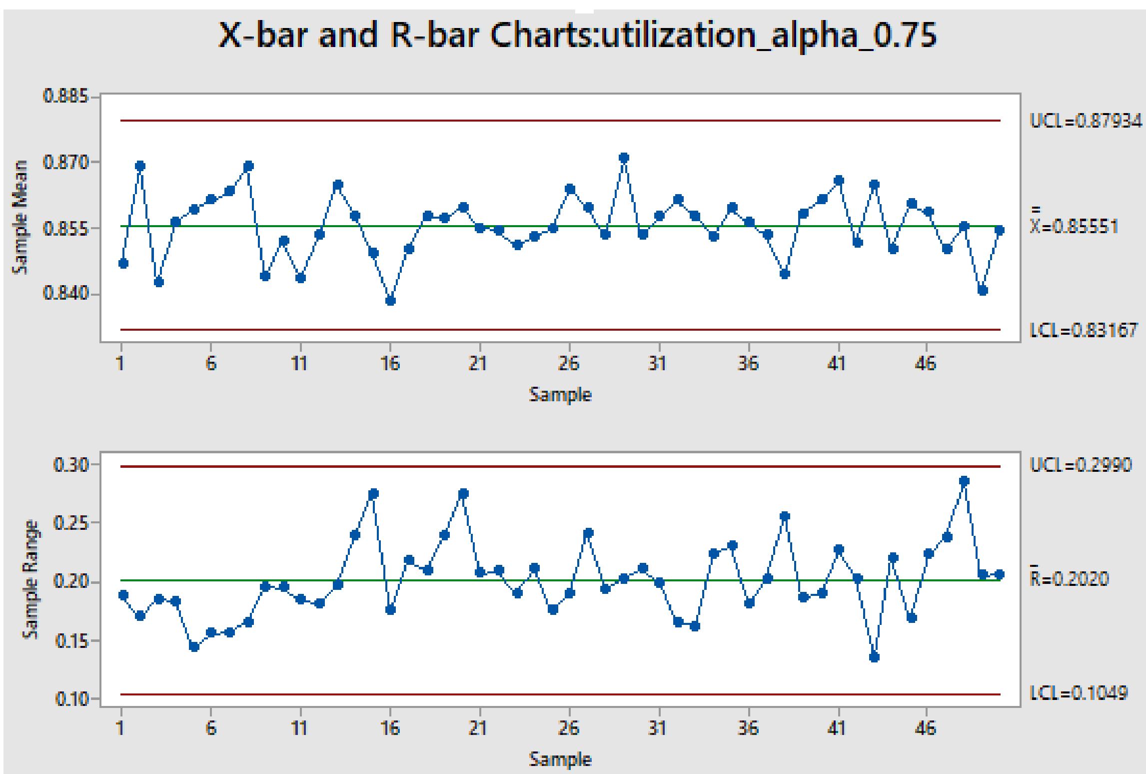

Figure 4. 21:X-bar and R-bar Charts of service level with alpha = 0.75

Individual X-bar chart of service level is presented in Figure 4. 22. It can be observed that the service level is clearly under control and is marginally greater than the service level with equal preference. X-bar and R-Chart is used with the continuous data collected in subgroup size of thirty-six. The Mean (X-Bar) of each subgroup for service level is charted on the top graph and the Range (R) of the subgroup of service level is charted on the bottom graph in Figure 4. 23.

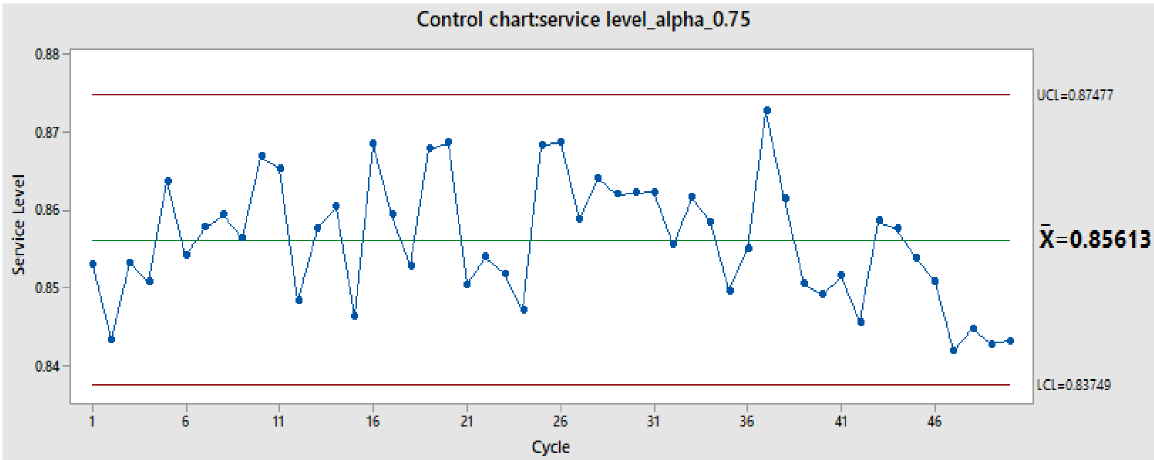

Figure 4. 22:Control chart of service level with alpha = 0.75

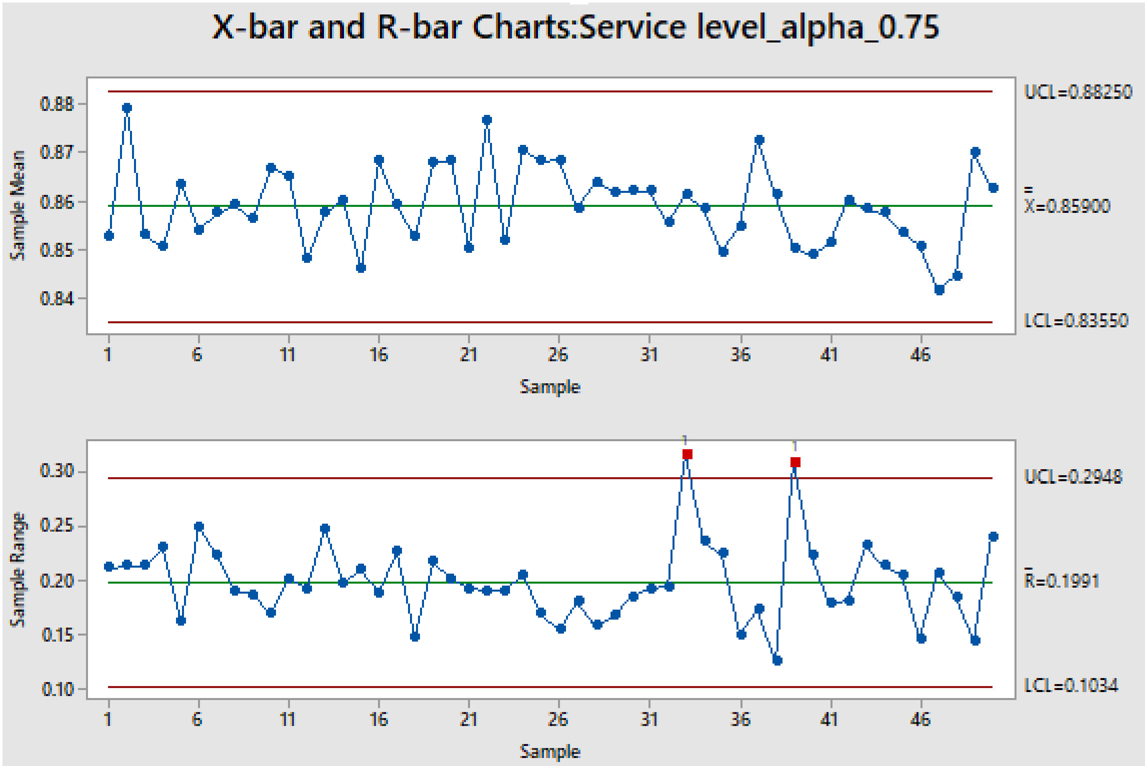

Figure 4. 23:X-bar and R-bar Charts of service level with alpha = 0.75

Statistical process control chart of trade-off is monitored along the time, as shown in Figure 4. 24, and is observed to clearly under the control with greater trade-off value than equal preference among the objectives.

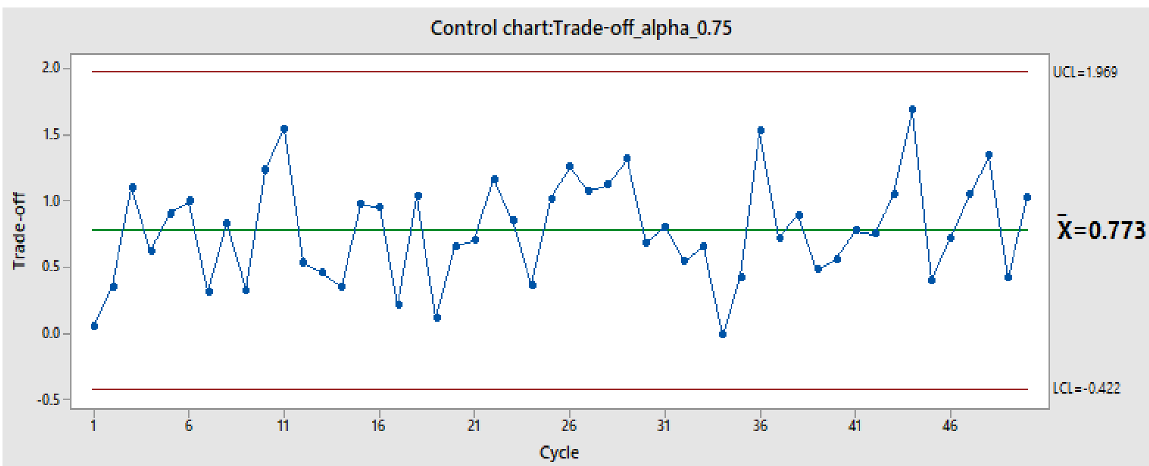

Figure 4. 24:Control chart of Trade-off with alpha = 0.75

Capability analysis is also carried out to verify, if the control limits of OR performance are within the specified limits. From Figure's 4.25 and 4.26, it is observed that process capability indices $c_p$and c_{pk} values are greater than one, indicating that the OR performance measures: utilization and service levels, are within the pre-established controllable limits.

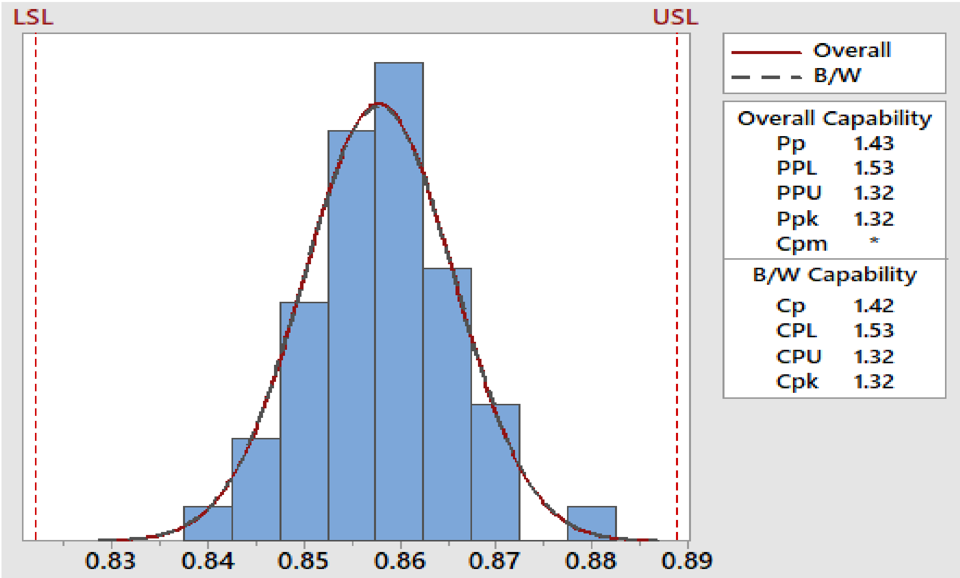

Figure 4. 25:Capability analysis of utilization with alpha = 0.75

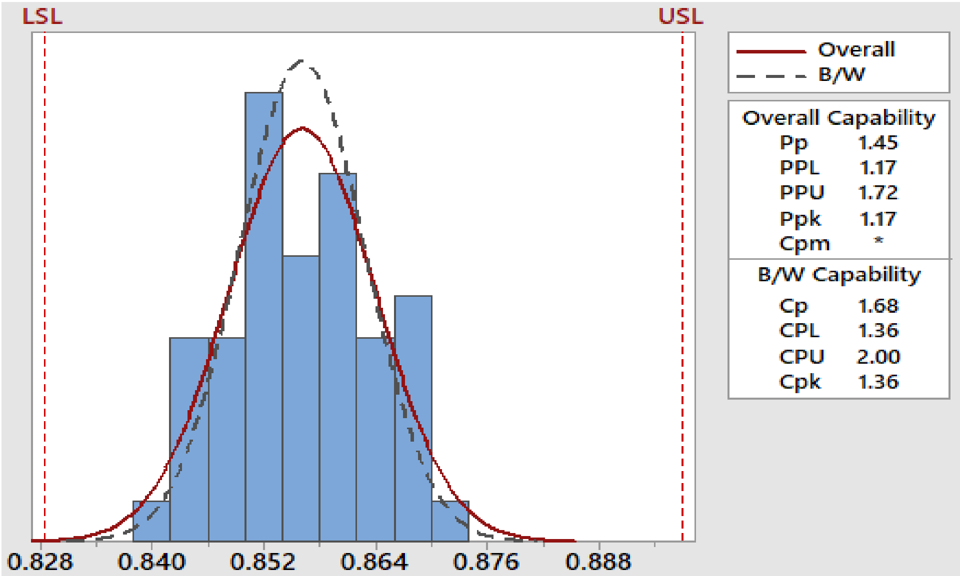

Figure 4. 26:Capability analysis of service level with alpha = 0.75

Control chart of preference adaptively changing along the cycles is monitored in Figure 4. 27. The mean (X-bar) of the preference is observed to be 0.7510 approximately equal 0.75, which is the preference chosen initially.

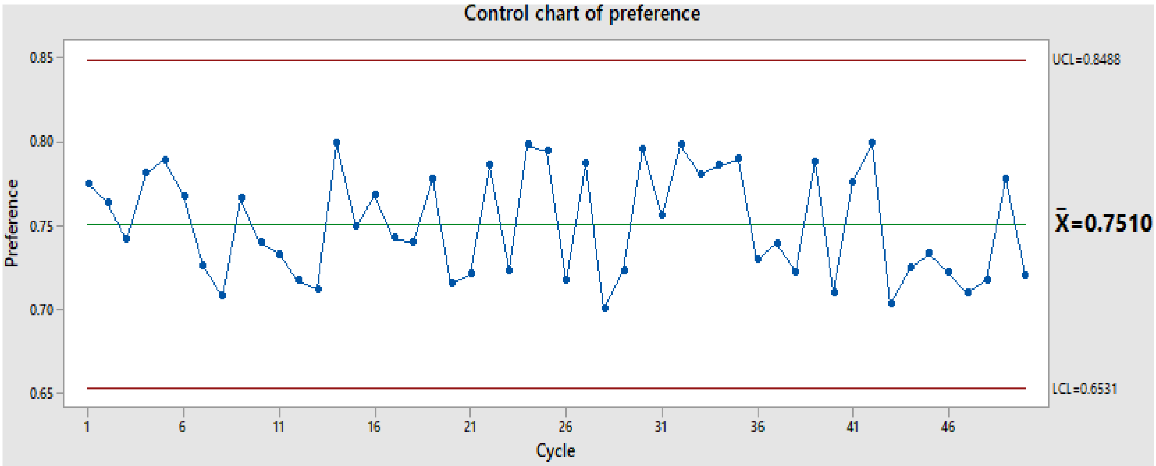

Figure 4. 27:Control chart of preference with alpha = 0.75

The promising results from the control charts and process capability analysis show that OR performance regarding utilization and service level are within the control limits along the time horizon. Therefore, it can be assured that our model of optimization and adaptive control enable OR managers to make an informed decision on OR configurations and control the OR performances in the long-term, given the stochastic demand and processing times of surgical specialties.

Chapter 5 Conclusion and future work

5.1 Conclusion

OR scheduling is essential, because of the rising demands, increasing expenditure and waiting times. ORs are liable for a significant proportion of admissions, therefore are the most work-intensive and cost consuming area of hospitals. High variations in patient arrival and processing times make the performance of OR plans highly unstable regarding utilization and service rate leading to high costs and long waiting lists. Current literature focusses on optimizing the OR configuration with predictive processing times and large inventory of demand, leading to low utilization and service level.

This thesis presents a three-step approach to optimize the OR configuration at strategic level planning. First, newsvendor model is used to balance the overtime and idletime costs in determining the block time and surgical-mix. Second, trade-off between the objectives, utilization and service level is exhibited and minimize using multiple-portfolio optimization. A one to one relationship is provided between distributions of demands, processing times, OR configuration and the distribution of expected OR utilization and service level. This relationship aides OR mangers in making an informed decision on OR configuration given stochastic demand and processing times of specialties. Third, an adaptive control scheme is proposed to ensure OR performance within predetermined control limits along the time horizon.

A simulation with normal distributions of demand and processing times of various surgical specialties at UKHC is used to validate the optimization model. Results demonstrate that the OR performance is well within the control limits along the time horizon.

5.2 Future work

This thesis is explicitly dealing with optimization of intra-operative stage, which is relatively very expensive when compared with the other stages of the peri-operative process. However, a holistic optimization of the peri-operative process will be a significant contribution towards achieving efficient healthcare system. Another significant direction in OR planning is to extend the optimization into other phases of planning down the line: tactical phase and Operational phase. This optimization model coupled with integer linear programming can be used to generate master schedule strategy within subspecialties. However, operational phase of OR planning needs a robust sequencing and scheduling methods to accommodate inherent variations in the system like cancellations, delays, no shows etc.…

References

Abedini, A., Ye, H., Li, W. (2016). Operating Room Planning under Surgery Type and Priority Constraints. Procedia Manufacturing, 5,15-25.

Abedini, A., Li, W., Ye, H. (2017). An optimization model for operating room scheduling to reduce blocking across the perioperative process. Procedia Manufacturing, 10, 60-70.

Adan, I. J. B. F., & Vissers, J. M. H. (2002). Surgical-mix optimisation in hospital admission planning: a case study. International journal of operations & production management, 22(4), 445-461.

Augusto, V., Xie, X., & Perdomo, V. (2010). Operating theatre scheduling with patient recovery in both operating rooms and recovery beds. Computers & Industrial Engineering, 58(2), 231-238.

Baligh, H. H., & Laughhunn, D. J. (1969). An economic and linear model of the hospital. Health services research, 4(4), 293.

Banditori, C., Cappanera, P., & Visintin, F. (2013). A combined optimization–simulation approach to the master surgical scheduling problem. IMA Journal of Management Mathematics, 24(2), 155-187.

Blake, J. T., & Carter, M. W. (2002). A goal programming approach to strategic resource allocation in acute care hospitals. European Journal of Operational Research, 140(3), 541-561.

Blake, J. T., Dexter, F., & Donald, J. (2002). Operating room managers' use of integer programming for assigning block time to surgical groups: a case study. *Anesthesia & Analgesia*, *94*(1), 143-148.

Cardoen, B., Demeulemeester, E., & Beliën, J. (2010). Operating room planning and scheduling: A literature review. European journal of operational research, 201(3), 921-932.

Dexter, F., & Ledolter, J. (2003). Managing risk and expected financial return from selective expansion of operating room capacity: Mean-variance analysis of a hospital's portfolio of surgeons. Anesthesia & Analgesia, 97(1), 190-195.

Dexter, F., & Traub, R. D. (2002). How to schedule elective surgical cases into specific operating rooms to maximize the efficiency of use of operating room time. Anesthesia & Analgesia, 94(4), 933-942.

Dexter, F., Epstein, R. H., Marcon, E., & Ledolter, J. (2005). Estimating the incidence of prolonged turnover times and delays by time of day. Anesthesiology: The Journal of the American Society of Anesthesiologists, 102(6), 1242-1248.

Dexter, F., Ledolter, J., & Wachtel, R. E. (2005). Tactical decision making for selective expansion of operating room resources incorporating financial criteria and uncertainty in subspecialties' future workloads. Anesthesia & Analgesia, 100(5), 1425-1432.

Ehrenfeld, J. M., Dexter, F., Rothman, B. S., Minton, B. S., Johnson, D., Sandberg, W. S., & Epstein, R. H. (2013). Lack of utility of a decision support system to mitigate delays in admission from the operating room to the post anesthesia care unit. Anesthesia & Analgesia, 117(6), 1444-1452.

Erdogan, S. A., & Denton, B. T. (2011). Surgery planning and scheduling. Wiley Encyclopedia of operations research and management science.

Forbes. 2015. U.S. Healthcare Spending on Track to Hit $10,000 Per Person This Year. http://www.forbes.com/sites/danmunro/2015/01/04/u-s-healthcare-spending-on-track-to-hit-10000-per-person-this-year/#68cb103b294c, (2016).

Guerriero, F., & Guido, R. (2011). Operational research in the management of the operating theatre: a survey. Health care management science, 14(1), 89-114.

Gupta, D. (2007). Surgical suites' operations management. Production and Operations Management, 16(6), 689-700.

Hans, E., Wullink, G., Van Houdenhoven, M., & Kazemier, G. (2008). Robust surgery loading. European Journal of Operational Research, 185(3), 1038-1050.

Kharraja, S., Albert, P., & Chaabane, S. (2006, October). Block scheduling: Toward a master surgical schedule. In Service Systems and Service Management, 2006 International Conference on (Vol. 1, pp. 429-435). IEEE.

Markowitz, H. (1952). Portfolio selection. The journal of finance, 7(1), 77-91.

Mulholland, M. W., Abrahamse, P., & Bahl, V. (2005). Linear programming to optimize performance in a department of surgery. Journal of the American College of Surgeons, 200(6), 861-868.

Olivares, M., Terwiesch, C., & Cassorla, L. (2008). Structural estimation of the newsvendor model: an application to reserving operating room time. Management Science, 54(1), 41-55.

Ozkarahan, I. (2000). Allocation of surgeries to operating rooms by goal programing. Journal of Medical Systems, 24(6), 339-378.

Porteus, E. L. (2002). Foundations of stochastic inventory theory. Stanford University Press.

Reddy Gunna, V., Abedini, A., & Li, W. (2017). Maximizing operating room performance using portfolio selection. Procedia Manufacturing, 10, 83-91.

Roberts, S., Saithna, A., & Bethune, R. (2015). Improving theatre efficiency and utilization through early identification of trauma patients and enhanced communication between teams. BMJ quality improvement reports, 4(1), u206641-w2670.

Santibáñez, P., Begen, M., & Atkins, D. (2007). Surgical block scheduling in a system of hospitals: an application to resource and wait list management in a British Columbia health authority. Health care management science, 10(3), 269-282.

Strum, D. P., Vargas, L. G., May, J. H., & Bashein, G. (1997). Surgical suite utilization and capacity planning: a minimal cost analysis model. Journal of Medical Systems, 21(5), 309-322.

Testi, A., Tanfani, E., & Torre, G. (2007). A three-phase approach for operating theatre schedules. Health Care Management Science, 10(2), 163-172.

Van Houdenhoven, M., van Oostrum, J. M., Hans, E. W., Wullink, G., & Kazemier, G. (2007). Improving operating room efficiency by applying bin-packing and portfolio techniques to surgical case scheduling. Anesthesia & Analgesia, 105(3), 707-714.

VanBerkel, P. T., & Blake, J. T. (2007). A comprehensive simulation for wait time reduction and capacity planning applied in general surgery. Health care management Science, 10(4), 373-385.

Viberg, N., Forsberg, B. C., Borowitz, M., & Molin, R. (2013). International comparisons of waiting times in health care–Limitations and prospects. Health Policy, 112(1), 53-61.

Vissers, J. M., Bertrand, J. W. M., & De Vries, G. (2001). A framework for production control in health care organizations. Production Planning & Control, 12(6), 591-604.

Wagner, H. M. (1969). Principles of operations research: with applications to managerial decisions. In Principles of operations research: with applications to managerial decisions. Prentice-Hall.

Wang, M. Y. (1999). Multiple-benchmark and multiple-portfolio optimization. Financial Analysts Journal, 55(1), 63-72.

Wang, Y., Tang, J., Pan, Z., & Yan, C. (2015). Particle swarm optimization-based planning and scheduling for a laminar-flow operating room with downstream resources. Soft Computing, 19(10), 2913-2926.

Zhang, B., Murali, P., Dessouky, M. M., & Belson, D. (2009). A mixed integer programming approach for allocating operating room capacity. Journal of the Operational Research Society, 60(5), 663-673.

Ye, H., Li, W., Abedini, A., Nault, B. (2017). An effective and efficient heuristic for no-wait flow shop production to minimize total completion time. Computers & Industrial Engineering, 108, 57-69.

Vita

NAME:

Vivek Reddy Gunna

EDUCATION:

Bachelor of Engineering 2015

Department of Mechanical Engineering

Jawaharlal Technological University Hyderabad, Hyderabad, India

PUBLICATIONS

Publications in referred international conferences

1. Reddy Gunna, V., Abedini, A. and Li, W., 2017. Maximizing operating room performance using portfolio selection. Procedia Manufacturing (*45th North American Manufacturing Research Conference*), 10, 83-91.

2. Reddy Gunna, V., Abedini, A. and Li, W., 2017. Multi-objective operating room planning for stochastic demand and case times. *Procedia Manufacturing* (*46th North American Manufacturing Research Conference*). (Submitted).