NURSE SCHEDULING AND RESCHEDULING PROBLEM UNDER UNCERTAINTY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL
OF
NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ECE KARPUZ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER SCIENCE
IN
INDUSTRIAL ENGINEERING

DECEMBER 2015

Approval of the thesis

NURSE SCHEDULING AND RESCHEDULING PROBLEM UNDER UNCERTAINTY

submitted by **ECE KARPUZ** in partial fulfillment of the requirements for the degree of **Master Science in Industrial Engineering Departments, Middle East Technical University** by,

Prof. Dr. Gülbin Dural Ünver
Dean, Graduate School of **Natural and Applied Sciences** __________________

Prof. Dr. Murat Köksalan
Head of Deparment**, Industrial Engineering** __________________

Asst. Prof. Dr. Sakine Batun
Supervisor, **Industrial Engineering Dept., METU** __________________

Examining Committee Members

Assoc. Prof. Dr. Sedef Meral
Industrial Engineering Dept., METU __________________

Asst. Prof. Dr. Sakine Batun
Industrial Engineering Dept., METU __________________

Asst. Prof. Dr. Melih Çelik
Industrial Engineering Dept., METU __________________

Asst. Prof. Dr. Mustafa Kemal Tural
Industrial Engineering Dept., METU __________________

Asst. Prof. Dr. Serhat Gül
Industrial Engineering Dept., TED University __________________

Date: 04.12.2015

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: ECE KARPUZ

Signature :

ABSTRACT

NURSE SCHEDULING AND RESCHEDULING PROBLEM UNDER UNCERTAINTY

Karpuz, Ece
M.S., Department of Industrial Engineering
Supervisor : Assist. Prof. Dr. Sakine Batun

December 2015, 86 pages

Nurse planning decisions play a critical role on hospital budgeting, quality of nursing services and nurse dissatisfaction. Nurse planning in a hospital includes four main phases which are nurse budgeting, nurse scheduling (rostering), nurse staffing (rescheduling) and nurse assignment. We consider the scheduling and rescheduling problems together under demand uncertainty. We formulate this problem as a two-stage stochastic integer program and consider different solution methods including solving the extensive form, L-shaped method and L-shaped based branch-and-cut method. To improve the efficiency of the decomposition methods, a lower bound is added and closed form of dual solutions of optimality sub problems are used while adding optimality cuts. Time series analysis is used to forecast the demand and nine months of historical data of Intensive Care Unit of a private healthcare provider is used for this purpose.

Keywords: Nurse Scheduling; Two Stage Stochastic Programming; L-Shaped Method; Time Series Analysis.

ÖZ

BELİRSİZLİK ALTINDA HEMŞİRE ÇİZELGELERİNİN OLUŞTURULMASI VE GÜNCELLENMESİ

Karpuz, Ece
Yüksek Lisans, Endüstri Mühendisliği Bölümü
Tez Yöneticisi: Yrd. Doç. Dr. Sakine Batun

Aralık 2015, 86 sayfa

Hemşire planlama, hastane bütçelemesinde, hastalara verilen hizmetin kalitesinde ve hemşirelerin iş memnuniyetinde kilit rol oynar. Bir hastanedeki hemşire planlama süreci temel olarak dört başlıktan oluşur; bütçeleme, çizelgeleme, çizelgelerin güncellenmesi ve hemşire atama. Biz bu çalışmada aylık çizelgelerin oluşturulması ve güncellenmesi problemlerini talep belirsizliği altında incelemekteyiz. Bu entegre problem iki aşamalı rassal tamsayılı programlama ile modellendi ve iki aşamalı modelin çözümünde extensive form, L-shaped algoritması ve branch ve cut'a dayalı L-shaped algoritması kullanıldı. Çözüm performansını geliştirmek için modele bir alt sınır kısıtı eklendi ve optimallik kısıtları alt problemin dual çözümünün kapalı formda yazılması vasıtasıyla bulundu. Talep tahmininde zaman serileri analizi kullanıldı. Analiz aşamasında özel bir hastanesinin yoğun bakım ünitesine ait 9 aylık veri kullanıldı.

Anahtar Kelimeler: Hemşire Çizelgeleme; İki Aşamalı Rassal Programlama; L-Shaped Metot; Zaman Serileri Analizi

ACKNOWLEDGEMENTS

I would like to thank my supervisor Asst. Prof. Dr. Sakine Batun for her support, guidance and insight throughout the research.

I would like to express my genuine gratitude to my family for their everlasting support and understanding.

I would also thank to my friends Ayşe Gül Irgıt, Devran Dönmez and Merve Aytekin for their support.

This work is supported by TUBITAK-BIDEB 2228 National - MSc and PhD Scholarship Programme for Senior Undergraduate Students.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

Memiş [1] states that annual healthcare expenditures of Turkey increased 13% on average between 2008 and 2011. This significant increase has directed hospital managements to explore reasons and to use their resources in more efficient ways. The key point is to reduce costs without sacrificing service quality. In order to provide this, efficient use of limited resources became inevitable and operations research methods have started to be used in hospitals.

Since nurses are one of the most important scarce resources in a hospital and nursing services have a big impact on both hospital budgeting and the quality of service provided, studying issues related to nursing services can help hospital managements to make progress.

Punnakitikashem [2] classifies the nurse planning decisions in a hospital in four categories. First step is nurse budgeting, which includes the long-term decisions such as the number of nurses to be hired and annual budget for nursing services. The second step is nurse scheduling or nurse rostering, in which the volume of patient arrivals is estimated and the assignment of nurses to shifts is made. Decisions made in this stage are referred to as mid-term decisions. The third step is nurse rescheduling, which includes short-term decisions such as rescheduling by making adjustments on the number of nurses available to meet the realized demand. The last step is nurse assignment, in which assignment of nurses to patients is made. In this study, we focus on nurse scheduling and nurse rescheduling decisions.

Making nurse scheduling and rescheduling decisions can be very critical as they have a direct impact on service quality, nurses' job satisfaction and hospital budgeting. Since these decisions are typically made by head nurses, it causes a high pressure on head nurses, takes too much of their time and requires a great deal of effort. Scheduling and rescheduling decisions are compelling because of the reasons given below:

- Demand is stochastic, there are variations in staffing requirements between days and even between shifts
- Maintaining an acceptable service level at all times is compulsory.
- Nursing services require qualified nurses.
- There are limited resources.
- Equity between nurses about their working times and satisfying their special requests are important.
- There are legal rules about working times, which are compulsory to be considered.

Building a poor schedule creates excess workload and high variability of daily workload on nurses, which decreases the quality of service provided by nurses.

A poor schedule also causes nurse dissatisfaction, which arises from:

- Excessive amount of changes on the schedule (high rate of calling on-call nurses, working overtime, etc.) during the month, which brings inconsistent working times for nurses.
- Not being able to meet special requests of nurses about working times.
- Not being able to provide equity between in terms of their working hours.

According to the interviews made by the head nurse, there is high nurse turn-over rate in the hospital, because of nurses' job dissatisfaction. Since qualification is so important in nurse services, training cost has an important role on budgeting and high turn-over rate causes high training costs.

We handle the scheduling and rescheduling problem together and consider the stochastic structure of demand. A two-stage stochastic programming model is presented in which the objective is to minimize adjustment actions during a month. The first stage decisions include mid-term decisions, which are monthly scheduling decisions and the second stage decisions include short-term decisions, which are adjustment (rescheduling) decisions those are made when the demand is known for certain. L-shaped method is used to solve the two-stage stochastic model. In order to improve efficiency, a lower bound is added and dual solutions of optimality sub problems are used while adding optimality cuts. Time series analysis is used to forecast the demand and nine months of historical data is included in the analysis.

CHAPTER 2

BACKGROUND INFORMATION AND PROBLEM DEFINITON

Although nurse scheduling/rescheduling is a problem encountered in all hospitals and shares common aspects across all types of hospitals (public, private, general, specialty, teaching, etc.), there are also hospital-specific or even ward-specific details of the problem since each system is unique. In this study, we particularly consider the intensive-care unit (ICU) nurse scheduling/rescheduling problem of a private hospital in Ankara, which has been in service since 2010. There are 25 active departments. The ICU includes 4 sub-units, which are Cardiovascular Surgery ICU, General ICU, Coronery Intensive Care Unit, Neonatal Intensive Care Unit. Our focus is on the nurse scheduling process in the Cardiovascular Surgery Intensive Care Unit and General Intensive Care Unit. There are 12 beds and 17 nurses in total. The total number of inpatients served during a month is 45 on average, and the average length of stay of a patient is 4.5 days.

According to the interviews made by the head nurse, there is a high turnover rate in nursing services and the reason is nurse dissatisfaction and undesirable schedules and overtime are the main reasons which cause dissatisfaction.

2.1 Scheduling

Scheduling process in the ICU includes the assignment of nurses to shifts and determination of off-days for each nurse.

There are 3 shifts in a day. The first shift includes the hours from 07:00 to 16:00, the second shift is from 15:00 to 24:00 and the third shift (night shift) is from 23:00 to 08:00.

There are four types of nurses:

- Scheduled nurses in a shift are the ones who are assigned to that shift in the monthly schedule.
- On-call nurses are also assigned in the monthly schedules and they should be prepared in case of calling when there is overload in the ICU.
- Overtime nurses in a shift are the nurses who are scheduled to work in that shift in the last minute (i.e., who are rescheduled to work) when the scheduled and on-call nurses are not enough to meet the workload.
- Undertime nurses in a shift are the nurses who are allowed to go when the actual demand is less than the planned demand and there is redundant workforce.

Monthly schedules are prepared manually by the head nurse of the ICU. The head nurse prepares the monthly schedule at the beginning of each month and makes these decisions:

- How many nurses will be assigned for each shift?
- Which nurse will work at which shift?
- Which nurse will be an on-call nurse at which shift?
- When are the off-days for each nurse?

A monthly schedule made by the head nurse at the beginning of month is given in Figure A.1 as an example.

Preparing these schedules manually takes significant time for head nurses and coming up with a desirable schedule can be very compelling. A desirable schedule should:

- meet legal working limit, rules and permissions,
- provide equity between nurses,
- be reasonably stable during the month,

- meet special requests of the nurses,
- result in a low rate of calling on-call nurses, overtime and undertime hours.

2.2 Rescheduling

The last hour of a shift overlaps with the first hour of the next shift. During this hour, situations of patients are told to the incoming team of nurses and the operations of next shift are determined. At the end of each shift, the charge nurse observes the actual demand for patient care and takes rescheduling decisions given below:

- If there is a shortage, the on-call nurse is called to work primarily.
- If on-call nurse is not enough to meet over workload, a nurse is assigned to work as an overtime nurse.
- If there is redundant workforce, the excess number of nurses will be allowed to go.

Because of the variability in demand, rescheduling decisions are made extensively. This sometimes causes excessive amount of changes in the existing schedule, high rate of calling on-call nurses and excess workload on nurses.

The monthly schedule given above as an example of monthly scheduling is also given in Figure A.2, but that representation is the schedule which includes the changes (i.e., rescheduling actions and disruptions) made during the month.

According to these schedules some observations are made:

- The schedules of nurses are changed during a month with an average of 8 days in a month for each nurse.
- Nurses are assigned as an on-call nurse on their days-off with a rate of 12% of days off (the number of assignments in which nurses are assigned as on-call nurses on their days-off / the total number of assignments as an on-call nurse).
- There are nurses who are assigned four consecutive night shifts.

- The maximum difference between the numbers of shifts in which nurses are assigned as an on-call nurse is 4 shifts.
- The rate of calling on-call nurses is 11% (The number of calling the on-call nurse / the total number of assignments as an on-call nurse).
- The total over-time need is 13% of total assigned normal workforce during the month.

2.3 Forecasting

In the current system the monthly schedules and rescheduling decisions are based on the experience of the head nurse. The demand for patient varies during a month and it can change even between consecutive shifts. Ignoring this variability is one of the main reasons which cause a poor schedule and end up with nurse dissatisfaction.

CHAPTER 3

LITERATURE REVIEW

We present the reviewed literature in two broad categories, which are review of methodologies and application based studies.

3.1. Review of Methodologies

This study basically consists of two parts, which are monthly scheduling under demand uncertainty and forecasting demand for nursing services. Stochastic programming and L-shaped method are used for modeling and solving the scheduling problem. In order to create the scenarios for stochastic programming by forecasting the demand, univariate time series analysis is used. A brief review of these methodological tools is presented below.

3.1.1 Stochastic Programming

Stochastic programming, which was first introduced by Dantzig [3], is mathematical programming where the problem parameters are random variables. As stated in Punnakitikashem [2], since real world problems typically include uncertainty, stochastic programming has a wide range of application areas including finance, manufacturing, transportation, logistics, airline operations, capacity planning and telecommunications.

A two-stage stochastic program is the simplest form of a stochastic program. In a two-stage stochastic program, decision variables are divided into two groups, which are first stage decision variables and second stage decision variables. First stage decision variables are the variables decided in the first stage before the actual realization of the random parameters. Second stage variables are the variables decided once the actual values of random parameters are realized.

The so-called two-stage stochastic program with recourse is of the form given in Birge and Louveaux [4] as:

$$\min c^T x + E_\varepsilon Q(x, e)$$

Subject to:

$$Ax = b$$

$$x \geq 0,$$

where $Q(x, e) = \min\{q^T y | Wy = h - Tx, y \geq 0\}$, ε is the vector formed by the components of q^T, h^T, and T, and E_ε represents the expectation with respect to ε. W is assumed to be fixed (fixed recourse).

As the number of scenarios in a stochastic program increases, solving the extensive form becomes computationally impractical. Therefore, using decomposition-based methods to solve stochastic programs is very common. One of such methods is the L-shaped method, which is basically the application of Benders decomposition to two-stage stochastic programs.

The main idea of the L-Shaped method is to approximate the expected second stage objective function value (i.e., the recourse function) by using a surrogate variable, Q, within an iterative framework. At each iteration, a restricted master problem (RMP) is solved to obtain a first-stage solution. If the solution is not feasible/optimal, a feasibility/optimality cut is added and next iteration is performed. Otherwise, the returned solution is the optimal solution.

The extensive form of the two stage stochastic programming model can be formulated as given below, in which K represents the all possible realizations, and p_k represents the probability of occurrence of the k^{th} realization:

$$\min c^T x + \sum_{k=1}^{K} p_k \, q_k^T y_k$$

Subject to:

$$Ax = b$$

$$T_k x + W y_k = h_k \qquad k = 1, \dots, K$$

$$x \geq 0, \qquad y_k \geq 0 \qquad\qquad k = 1, \dots, K$$

Because of the block structure of extensive form, the following algorithm is named as L-shaped method. Birge and Louveaux [4] state that this structure makes a Benders decomposition or equivalently a Dantzig-Wolfe decomposition of its dual possible. This method has been extended in stochastic programming to take care of the feasibility issue and is known as the L-shaped method. It proceeds as follows:

Step 0: Set $r = s = v = 0$.

Step 1: Set $v = v + 1$. Solve the following LP:

$$\min \text{z} = c^T x + \theta \tag{1}$$

Subject to:

$$Ax = b$$

$$D_l x \geq d_l \qquad l = 1, \dots, r \tag{2}$$

$$E_l x + \theta \geq e_l \qquad l = 1, \dots, s \tag{3}$$

$$x \geq 0, \qquad \theta \in \mathbb{R}$$

Let (x^v, θ^v) be an optimal solution. If no constraint (3) is present, θ^v is set equal to (-∞) and is not considered in the computation of x^v.

Step 2: For $k = 1, \dots, K$ solve the following LP:

$$\min \text{w}' = e^T v^+ + e^T v^- \ (4)$$

<u>*Subject to:*</u>

$$Wy + Iv^+ - Iv^- = h_k - T_k x^v (5)$$

$$y \geq 0, \quad v^+ \geq 0, \quad v^- \geq 0$$

And let σ^v be the associated simplex multipliers (i.e., dual variables). If $w' > 0$, define

$$D_{r+1} = (\sigma^v)^T T_k$$

and

$$d_{r+1} = (\sigma^v)^T h_k$$

to generate a constraint (called a feasibility cut) of type (2). Set $r = r + 1$, add to the constraint set (2), and return to Step 1. If for all k, $w' = 0$,go to Step 3.

Step 3: For $k = 1, \dots, K$ solve the following LP:

$$\min \text{w} = {q_k}^T y \qquad (6)$$

<u>*Subject to:*</u>

$$Wy = h_k - T_k x^v$$

$$y \geq 0$$

Let ${\pi_k}^v$ be the simplex multipliers (i.e., dual variables) associated with the optimal solution of Problem k of type (6). Define

$$E_{s+1} = \sum_{k=1}^{K} p_k * ({\pi_k}^v)^T y_k \, T_k$$

and

$$e_{s+1} = \sum_{k=1}^{K} p_k * (\pi_k{}^v)^T y_k \, h_k$$

Let $w^v = e_{s+1} - E_{s+1}x^v$. If $\theta^v \geq w^v$, stop; $x^v$is an optimal solution. Otherwise, set $s = s + 1$, add to the constraint set (3) and return to Step 1 [4,5].

There are some measures used to evaluate the impact of uncertainty.

The expected value of perfect information ($EVPI$) measures the amount of payment that a decision maker is willing to pay in return for complete and accurate information about future. In order to determine $EVPI$, firstly one needs to solve a deterministic model for each realization (ε) of the random variables, then find the expected value of optimal objective values of these solutions. This is called the wait-and-see solution (WS). $EVPI$ is obtained by comparing the wait-and-see solution to the here-and-now solution corresponding to the recourse problem (RP); i.e., stochastic program:

$$EVPI = RP - WS$$

It is assumed that for all ε, there exists at least one feasible solution (which implies there is at least one optimal solution), otherwise there is no chance to construct a reasonable stochastic model.

A heuristic solution of the model with random parameters can be obtained by replacing random variables with their expected values and solving a deterministic model. This approach is called as the expected value or mean value problem (EV). The value of the stochastic solution (VSS) is defined as the possible gain obtained when the stochastic model is solved. VSS is calculated as given below:

$$VSS = EEV - RP$$

where EEV is the expected value of the mean value solution under each realization ε [4].

3.1.2 Time Series Analysis

Time series analysis deals with analyzing and modeling an ordered sequence of observations which are generally obtained over equal time increments [6]. Compared to regression analysis, time series analysis has an advantage of taking the internal structure of data into consideration. To illustrate, the assumption about serial uncorrelated residuals is often violated in regression analysis. Taking this autocorrelation into account provides more realistic forecasts.

The assumption of time series analysis is that data are stationary, which requires the property that the mean, variance and autocorrelation structure do not change over time [7].

Dickey-Fuller is one of the methods for testing the existence of unit root (the situation of nonstationarity) in a series. Address the model given below:

$$Y_t = p * Y_{t-1} + u_t$$

Where u_t is the stochastic error term and p is the coefficient. We can show the equality as:

$$Y_t - Y_{t-1} = (p - 1) * Y_{t-1} + u_t$$

$$\nabla Y_t = \gamma * Y_{t-1} + u_t$$

where $\gamma = (p - 1)$. The main goal is testing the following null hypothesis:

H_0: $\gamma = 0$ (data contains the unit root)

H_1: data is stationary.

If $|\gamma + 1| < 1$, then data is said to be stationary. Testing is made by t statistics [7].

Akaike's Information Criteria (AIC) and Bayesian Information Criteria (BIC), which is also known as Schwarz Information Criteria (SIC), are the two methods used commonly in order to select the best model that represents the data. The goal of AIC is to find the best approximating model to the unknown true data generating process. AIC selects the model that minimizes the negative likelihood penalized by the number of model parameters.

$$AIC = -2\, logp(L) + 2m$$

where L represents the likelihood under the fitted model, m represents the number of parameters in the model and p is the number of lags.

The aim of BIC is to find the most probable model. BIC is given as:

$$BIC = -2\, logp(L) + plog(n)$$

The difference between the AIC and BIC representation is BIC depends on the sample size n [8].

The ideal case for the best model is the selecting the model with the minimum AIC and SIC values.

The autoregressive (AR) model is the common approach for modeling univariate time series:

$$X_t = \delta + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + A_t$$

where X_t is the time series, A_t is white noise, $\phi_1, \dots, \phi_p$ are the coefficients of the model, and

$$\delta = (1 - \sum_{i=1}^{p} \phi_i)\mu$$

with μ denoting the process mean [9].

A white noise process is one with no discernible structure. A definition of a white noise process is given in Rachev et al. [10] as:

$$E(y_t) = \mu$$

$$var(y_t) = \sigma^2$$

$$\gamma_{t-r} = \begin{cases} \sigma^2 & if\ t = r \\ 0 & otherwise \end{cases}$$

Therefore a white noise process has constant mean and variance, and zero autocovariances.

An autoregressive model is said to be simply a linear regression of the current value against prior value, in other words the current value of the series depends on only the values that occur in previous time periods and error term. The value of p is called the order of the AR model [7, 9].

The moving average (MA) model is another common approach for modeling univariate time series and is represented as:

$$X_t = \mu + A_t - \theta_1 A_{t-1} - \theta_2 A_{t-2} - \cdots - \theta_q A_{t-q}$$

where X_t is the time series, μ is the mean of the series, A_{t-i} are white noise terms, and $\theta_1, \ldots, \theta_q$are the parameters of the model. The value of q is called the order of the MA model. A moving average model is said to be a linear regression of the current value of the series against the white noise of one or more prior values [9].

Box-Jenkins ARMA (Autoregressive Moving Average) model is the model where AR and MA models are used at the same time. ARMA(p, q) model is obtained by combining AR(p) and MA(q) models is represented as follows:

$$X_t = \delta + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + A_t - \theta_1 A_{t-1} - \theta_2 A_{t-2} - \cdots - \theta_q A_{t-q}$$

It can be inferred that in ARMA models the current value of series depends linearly on its previous values and a combination of current and previous values of a white noise error term [10].

The Box-Jenkins models assume that time series is stationary and in case of an non-stationary situation, stationarity can be obtained by differencing non-stationary data one or more times. When differencing is applied, ARMA model turns to be ARIMA model in which “I” represents the “Integrated” term and next step will be identifying the order of AR and MA models which are also represented with p and q also. Autocorrelation and partial autocorrelation functions (ACF and PACF) are the primary tools for identifying the order of AR and MA models [9].

The autocorrelation function (ACF) shows how the value is correlated to previous values; more specifically autocorrelation in lag k, called p_k, is simply the correlation between the values from p_t to p_{t-k} for stationary processes. Autocorrelation function shows the randomness in data. As a proof of randomness, autocorrelations are expected to be near zero for any and all time-lag separations. If one or more autocorrelations are significantly non-zero, then data is said to be non-random. The partial auto correlation function (PACF) measures the correlation only between an observation k periods ago and the current observation, after controlling for observations at intermediate lags. Partial autocorrelations are useful in identifying the order of an autoregressive model [9, 10].

The differences in ACF and PACF among models are useful when selecting models. ACF is used to identify MA models and PACF is used to identify AR models. The following Table 3.1 summarizes the ACF and PACF behavior for these models [11].

Table 3.1 ACF and PACF Behavior for AR, MA and ARMA Models

Conditional Mean Model	ACF	PACF
AR(p)	Gradually decrease	Cuts off after p lags
MA(q)	Cuts off after q lags	Gradually decrease
ARMA(p,q)	Gradually decrease and cuts of after q lags.	Gradually decrease and cuts of after p lags.

Decrease can be exponential or sinusoidal wave [12].

To sum up the basic properties of ARIMA models are given below as mentioned in Weggemans [13]:

- Non-stationary data can be made stationary by taking differences of the original series.
- The residuals of the estimated ARIMA series should follow a normal distribution and should not possess autocorrelation.
- The Akaike criterion can be used to identify the best among the estimated ARIMA series.

3.2 Application Based Studies

Application based studies consist of five main topics which are employee scheduling, deterministic and stochastic nurse scheduling, patient volume forecasting, measuring nurse workload.

3.2.1 Employee Scheduling

The studies about employee scheduling include days-off scheduling and generating working plans under some uncertainties. Employee scheduling studies are similar to nurse scheduling studies in terms of dealing with uncertainties and days off scheduling.

Therefore the studies given below are included in the literature review since nurse scheduling includes assignment of days-off to nurses and deals with similar uncertainties.

Morton and Popova [14] study an employee scheduling problem where the demand forecast of the required number of shafts per type with due dates for the next month is received at the end of each month. One of the line manager's tasks is to build an employee schedule for shaft production for the next month. Machines have different production rates for each shaft type and for each crew as well as different down-time rates. The production line manager decides which work will be assigned to which crew and at which machine they will work to meet the required number, which is forecasted on time and within the budget constraint. In this problem, the random parameters are production rates and machine availabilities. To maintain this randomness, Bayesian distributional forecast is used. The distributions are updated with observations of each passing month. Monthly employee scheduling is made by solving a two stage stochastic program with recourse. The Bayesian estimation model provides point and distributional estimates for the hourly production rates by shift and shaft type and for up times of the production equipment. These estimates are used as inputs for the deterministic optimization model first. This model minimizes a weighted sum of penalties for late and non-delivered shipments plus a penalty for exceeding the target budget. Since this problem considers production rates and down-time rates as known certainly, deterministic model is extended to a stochastic model. Morton and Popova [14] state that Bayesian forecasting models can rapidly capture changes in non-stationary systems using limited historical data.

Alfares [15] studies employee days-off scheduling in which work/off days are determined for a work week. In this problem, daily labour demands are random variables and a simulation model is used. The relevant unit has 19 employees divided into five craft types and employees can be assigned to three types of days-off schedule. The aim of the study is to find technician's days-off schedules and meet labour demands by considering limited staff availability and policy restrictions on the choice of employee schedules. The number of technicians of each craft to assign to each days-off schedule is the decision that must be made while minimizing the average throughput (waiting plus processing) time of maintenance work orders (W/O). The number of required technicians of each craft varies from one work order to another and the number of technicians from each craft type assigned to W/Os is calculated from empirical probability distributions. Some W/Os need more than one craft type therefore historical data is used to determine the percentage of time each craft is needed by a given W/O. It is assumed that employees work at an average speed and they are fully available during the simulation period.

Campbell [16] investigates the employee days off scheduling with random demand. A two stage stochastic programming model is built. In the first stage, days-off scheduling is made and in the second stage, assignment of cross-trained workers is made to meet actual demand.

3.2.2 Nurse Scheduling

Nurse scheduling problems are widely studied in the operations research literature. The problem is to develop a decision making tool that assigns nurses based on nurse preference and patient workload requirements. Nurse scheduling literature includes a wide range of studies. These studies can be classified in various ways. Studies based on nurse scheduling in operating suites and nurse scheduling in general clinics are the most common subjects in nurse scheduling literature. Some studies handle a single-objective and generally the objective in these studies is either minimizing costs or maximizing nurse preferences. On the other hand, some studies consider multiple-objective optimization problems. In these studies, the objective function is usually minimizing the total penalty cost that occurs due to the violations of soft constraints.

Literature about nurse scheduling can also be classified as cyclic and non-cyclic scheduling. In cyclic scheduling, a predetermined working pattern which is repeated in every scheduling period is assigned to each nurse. In non-cyclic scheduling, a new schedule is made at the beginning of each scheduling period.

Commonly used solution approaches for the nurse scheduling are solution methods based on mathematical programming and heuristic methods.

Most studies assume that decision makers have complete information and handle nurse scheduling problem as a deterministic problem. However, healthcare organizations deal with many different types of uncertainties and considering these uncertainties is critical when modeling and solving the planning problems in healthcare delivery systems. New efforts involve forecasting the staff requirement for the near future. One of the key factors of high quality is assigning the correct number of personnel to meet the requirement [17, 18, 19].

Kao and Queyranne [20] introduce eight models including single period/multi period, aggregate/disaggregate and deterministic/probabilistic model. In a single period model, the time-varying nature of demand for nursing hours is ignored. Aggregation is done over the nurse skill classes. In probabilistic models, demand uncertainty is considered. It is indicated in the study that ignoring the time-varying nature of demand does not cause gross errors in budget estimates, on the other hand ignoring demand uncertainty induces error about five to six percent on budget estimates.

In order to investigate general structure of nurse scheduling model, the deterministic nurse scheduling studies are analyzed. On the other hand, in practice, the nurse scheduling problem is a stochastic problem where demand is uncertain. As a result, studies about stochastic nurse scheduling and rescheduling are also included in the literature review. In addition to these studies, studies about forecasting patient volume and measurement of workloads are reviewed for this study, since forecasting the demand is one of the main tasks to be performed when scheduling the nurses under demand uncertainty.

Tein and Ramli [21] review the recent advancements on nurse scheduling. Nurse scheduling and rescheduling reviews can also be found in Cheang et. al. [19], Clark et. al. [22], Ernst [23].

3.2.2.1 Deterministic Nurse Scheduling

Belien and Demeulemeester [24] consider nurse scheduling and surgery scheduling at the same time. It is stated that a common problem at hospitals is the extreme variation in daily workload pressure for nurses and one of the main reasons for this variety is the operating room schedules. Therefore the study aims to save staffing cost by integrating operating room scheduling and nurse scheduling problems. The objective function of the presented model is to minimize the total required number of nurses. The workload distribution is the input for the nurse scheduling model. Constraints of the model consist of two main groups. One of them is the coverage constraints, which represent how many nurses of appropriate skills need to be scheduled for each demand period. The other one is collective agreement requirements, which are the constraints that define rules for an acceptable schedule in terms of workload, day-off and resting time between shifts. Instead of assuming the demand values which are the right-hand side values of the coverage constraints in the nurse scheduling problem are fixed, general nurse scheduling problem (GNSP) is studied. In GNSP, demand values are considered to be dependent on the workload patterns which will be obtained by enumerating all possible ways of assigning operating blocks to the different surgeons, subject to surgery demand and to capacity restrictions. Column generation technique approach is used to solve the IP.

Mobasher et al. [25] study daily scheduling of nurses in operating suites. They work with a variety of objectives such as minimizing the maximum demand deviation for any case, maximum amount of overtime assigned to any nurse, maximum number of cases assigned to any nurse and aim at Pareto-optimal solutions. A multi-objective integer program is used to formulate the problem. The aim is to determine which nurse should be assigned to which surgery case, during which time intervals, and what their role (Nurses can have different roles according to their skill level during a surgery, i.e., a scrub is responsible for preparing and passing supplies, equipments and instruments to the surgeon during the procedure) should be. The surgery durations and the number of required nurses in each role are the main parameters given to the model. Two methodologies are used to find a solution. The first one, which is called the solution pool method (SPM), generates a pool of good solutions by solving multiple optimization problems, each of which optimizes a single objective. Then a cumulative weighted index is found for each solution and the solution with the smallest index is picked as the best solution to be used. The second method is called the modified goal programming method (MGPM) and it finds the optimal solution for each goal separately and then solves a derivative optimization problem whose objective is to minimize the sum of the deviations from those goals. It is stated that although MGPM generates solutions with smaller deviations in significantly less time, SPM has the advantage of providing good solutions among which the decision maker can choose.

Bard and Purnomo [26] study the rescheduling problem, which aims to reallocate the available resources in a way that the cost of the shortfall is minimized while ensuring that each unit in the hospital has sufficient coverage. Decisions made in the integer programming model include overtime, outside nurses and floaters. In doing so, minimizing the differences between the new plan and the original plan is also considered and the expected demand for the upcoming 24 hours is taken as an input.

Glass and Knight [27] state that a nurse rostering problem includes two constraint types. Staffing constraints ensure that sufficient nurses of each type are on duty at any particular time and schedule constraints are related to the sequences and combinations of shifts to be worked by each nurse. It is said that satisfying both sets of constraints simultaneously is not always possible. The modeling approach therefore involves reducing selected constraints to soft constraints with measurements of their violation. The objective is then to minimize the violation of these soft constraints. MIP is used for modeling and a methodology for handling continuity between rostering periods is studied.

Atmaca et al. [28] study nurse scheduling problem and use 0-1 linear programming to formulate the problem. There are three objectives considered, which are minimizing the total number of working days of nurses, minimizing the difference between the total numbers of working days of nurses and minimizing the number of assignments of nurses to consecutive shifts. The objective function is represented as the minimization of the weighted sum of the deviations from these goals related to each objective.

3.2.2.2 Stochastic Nurse Scheduling

Punnakitikashem et al. [29] model the nurse assignment problem under uncertainty in the workload as a two stage stochastic integer program. Since a patient may be admitted or discharged during a shift, the amount of direct care required by the patient may vary dramatically throughout the shift. The first stage decision is assigning nurses to patients and the second stage decision is determination of realized workload. The aim is to minimize excess workload on nurses. Benders' decomposition method, in which the master problem assigns nurses to patients, and each recourse problem penalizes the assigned workload is used as the solution approach. The proposed approach decomposes by scenario and also by nurse into the $(number\ of\ scenarios) * (number\ of\ nurses)$ linear programming subproblems. Therefore, the subproblems become more manageable than subproblems decomposed by standard L-shaped method.

Punnakitikashem [2] builds a two stage stochastic programming model where nurse staffing and nurse assignments are integrated. Workload on the nurses is uncertain and the aim is minimizing excess workload on the nurses under a budget constraint. The first stage decision is to assign nurses to patients and the second stage includes rescheduling decisions in which the decision of assigning overtime nurses, agency nurses or cancelling scheduled nurses. Three solution approaches; namely Benders

Decomposition, Lagrangian Relaxation with Benders Decomposition, and Nested Benders Decomposition are presented. Firstly, Benders Decomposition is used to solve two stage stochastic model, then Lagrangian Relaxation with Benders Decomposition is used to solve model in which the budget constraint is relaxed. Secondly, the problem is considered as a multistage stochastic programming problem and Nested Benders Decomposition is demonstrated. An algorithm for finding non-dominated solutions obtained from these three approaches is presented. Non-dominated solution is defined as nurse schedules and assignments that are not dominated by any other schedules and assignments found, either they require less excess workload or less staffing cost than the other solutions found.

Kim [30] builds an integrated staffing and scheduling (iStaff) model as a two-stage stochastic integer program with mixed integer recourse. Demand is uncertain. Staffing decisions are made well ahead in time and when the demand is known for certain, adjustments are made. As a result, the first stage decision is the determination of the number of nurses who will work in pregenerated scheduling patterns at any time and the second stage decisions are adjustment decisions including amount of overstaffing, amount of understaffing, etc. It is stated that the problem size is large because staffing and scheduling decisions include a high number of integer variables because of the possible shift combinations. L-shaped method is used to solve the model. The major contribution of this paper is defined as identifying valid mixed integer rounding (MIR) of feasible solutions for the second stage mixed integer programming problem and exploring heuristic approaches for cut aggregation strategies and branching strategies tailored to the model formulation.

In most of the previous studies, the stochastic structure of demand is ignored. Punnakitikashem et al. [29] consider only the nurse assignment under uncertainty on the amount of direct care required by the patient. In the next study, Punnakitikashem et [2], the model in Punnakitikashem et al. [29] is extended by incorporating the nurse staffing decisions into the assignment model. Since a short-term nurse staffing is considered, nurse preferences are not included in the model, which are normally included in mid-term scheduling. There are a limited number of studies that integrate nurse scheduling and rescheduling. Kim [30] uses the scheduling patterns which are pregenerated to ensure compliances with scheduling rules and regulations and in the first stage determines the number of nurses who will work in these scheduling patterns. This method disregards the special requests and special occasions of nurses (like breast-feeding permissions). Our model provides more flexible schedules in terms of these conditions. The demand uncertainty is handled via patient volume data. Since the required care by a patient is different among patients, the indicator that shows the total required patient care in terms of all patients is used to forecast demand in our study.

3.2.3 Forecasting of Patient Volume

Weggemans [13] handles three issues, namely building a model to predict the number of patient arrivals in a certain time period, a model that can compute the probability that a patient will transfer from one specialism to another in a certain time period and a model that can estimate the service time of patients. Markov chains are used to model the transition probabilities. Some studies that use time series to predict the patient volumes are referenced. It is stated that using queuing models is not an efficient way to predict patient volumes because queuing theory requires a specific arrival distribution and service time distribution. As a result Autoregressive Integrated Moving Average (ARIMA) models are commonly used for prediction of patient volumes.

Schweigler et. al. [31] investigate how time series-based models perform in short-term forecasting of emergency department (ED) crowd. While patient arrivals per hour is used generally in most studies, in this study ED hourly occupancy, which is defined as the total number of patients (patients in each adult ED + patients in waiting room) divided by the number of permanent beds during that hour, is used. Three models, which are hourly historical average, seasonal autoregressive integrated moving average (SARIMA) and sinusoidal with an autoregression (AR) - structured error term is used for prediction. Comparison of these models is made according to log likelihood and AIC and accuracy of models are measured by actual observed bed occupancy with root mean square (RMS) error. Results show that while AR based models are not different from each other, they perform better compared to historical average model.

Jones et. al. [32] investigate the use of seasonal autoregressive integrated moving average, time series regression, exponential smoothing, and artificial neural network models to forecast daily patient volumes at three different facilities. Forecasts are made for horizons ranging from 1-30 days in advance. Accuracy of models is evaluated according to mean absolute prediction error (MAPE). The seasonal and weekly pattern of daily patient volume in ED services is confirmed. It is stated that the existing methodology proposed in the literature, multiple linear regression based on calendar variables, is an acceptable approach, on the other hand regression-based models that incorporate calendar variables, account for site-specific special-day effects, and allow for residual autocorrelation provide a more appropriate, informative, and consistently accurate approach to forecasting daily ED patient volumes.

Kam et. al. [33] evaluate three models, namely moving average, seasonal ARIMA (SARIMA) and multivariate SARIMA to predict the number of patients visit an emergency center per day. Three models are investigated by considering calendar and weather data. Residual analysis, ACI and Bayesian information criterion are used to compare goodness of fit. Accuracy of models are measured by MAPE. It is stated that the most appropriate and accurate model for predicting number of patients visiting ED is the multivariate SARIMA model.

Kao and Tung [34] state that there are two different approaches in forecasting a time series, which are the time series approach and the econometric approach. It also stated that although econometric approach generally gives better forecasts, it requires a large amount of data set. As a result it is stated that time-series models have been used to forecast patient census. In this paper ARIMA (Autoregressive integrated moving average) time series models for forecasting demands for inpatient services is studied. Prediction of demand is made on a yearly basis. Demand is stated in terms of monthly admissions and patient days by services. First, monthly admissions and patient days are forecasted by using ARIMA, and then the actual demand and forecasted demand are compared empirically. Finally an indirect method to project patient days is introduced. This approach combines admission forecasts and the length of stay estimates and takes less effort.

Kim [30] compares forecasting methods used while predicting hospital patient volume. Exponential smoothing model, ARIMA model, autoregressive moving average with generalized autoregressive conditional heteroskedasticity (ARMA-GARCH) model and vector autoregressive (VAR) model were investigated. Comparison is made according to MAPE. It is noted that the multivariate forecasting method used accounts for patient admissions to Hospital Medicine (HM) from a variety of sources (e.g. emergency medicine, outpatient offices, intensive care services, etc.), while the univariate methods use HM patient volume data only. The results show that a univariate ARIMA model performs best. It is stated the multivariate model does not perform better than the univariate ARIMA models, particularly for the forecast periods of more than five days.

3.2.4 Nurse Workload Measurements

Nurses take a critical role on the quality of healthcare system. They are the major factor that affects the quality of patient care. Continuity and strict care is inevitable and critical for especially ICU patients. The excess workload on nurses and the extreme turn-over rate which arises from dissatisfaction of nurses causes low quality of patient care. Therefore measuring workload on nurses accurately is an essential issue for hospital managers.

Nursing workload measurement systems are used to determine the amount of care needed by patients and required nursing time to meet those needs. In addition to this nursing workload measurement provides information to predict number of nurses required for next shift. Nursing workload measurement data can be a base for budgeting, staffing, planning decisions and quality assurance.

Some of nursing workload measurement methodologies are based on tasks nurses perform during a shift. Each task has a standard completion time and total time required for performing tasks gives the total nursing workload. On the other hand, some methodologies use patient classification systems. In consideration of some specific and predetermined factors, needs of each patient are evaluated and according to needs and features of patients each patient is assigned to a predetermined patient type [35].

In the 1960s a method called "Utilized Work Sampling" was used. This method does not consider the patient type and nursing skills. From the mid-1970s to 1990s a method which takes into account the severity of patient and the dependency degree of patient to nurses. In 1990s measurements based on performance of nurses became important due to pressure on willing to decrease nursing expenses as a result nursing ratio concept was started to be carried out.

Methods used currently to decide the required number of nurses can be classified under 5 categories [36]. These are;

- Professional Judgement Approach: This approach is based on the calculation of working hours of nurses decided in the schedules. When the qualification of nurses and the dependence level of patients change, this approach will be inefficient.
- Nurse Per Occupied Bed Method: This approach is based on the beds occupancy rate and ignores the levels of dependence of patients.

- Time-Task/Activity Approaches: The operations performed by nurses and the required time for these operations are recorded and analysis is made according to these records.
- Regression-based Systems: Regression analysis is made to forecast the required number of nurses for a specific task. Generally, the relation between the required number of nurses and the bed occupancy rate is analyzed.
- Acuity-Quality Method: In this approach, in addition to the number of patients, the level of dependence of each patient is also considered. This approach is efficient in the systems where the number of patients and the types of patients are changeable.

Padilha et. al. [37] underline the importance of measuring nursing workload and taking into account the indicators of workload to increase quality and safety of care given in hospital departments. They consider this necessity under “Nursing Activities Score (NAS)”.

Lin et. al. [38] focuses on quantitative models of work related fatigue. Two methods which are called survey-based and total function-based fatigue models are introduced. A multi-objective MIP formulation is used to model the scheduling problem. Objective function of the model is a weighted sum of total preference scores, total based survey-based fatigue scores and total function-based fatigue scores for all nurses at the end of their shift patterns. It is stated that making Pareto-optimal schedules is possible where the nurse fatigue levels are significantly reduced for a small decrement in nurse preferences.

3.2.4.1 TISS (Therapeutic Intervention Scoring System)

In the current system of the hospital, workload measurement is based on nurses per occupied bed method. The data used in studies is collected from TISS method, which is a task based approach.

There are four categories of tasks performed. Each category has a score from 1 to 4. For each patient tasks need to be performed during a day is checked on a check list via using Excel worksheet. Each task has its own coefficient. At the end of the checklist a total TISS score is obtained for each patient. According to this scoring system a nurse should have workload corresponds to 50 TISS score on average.

CHAPTER 4

MATHEMATICAL FORMULATION AND SOLUTION APPROACHES

4.1 Integrated Nurse Scheduling and Rescheduling Model

We formulate the problem as a two-stage stochastic programming model. In the first stage, mid-term (monthly) scheduling decisions are made:

- Which nurse will work at which shift?
- Which nurse will be an on-call nurse at which shift?
- When are the days-off for each nurse?

In the second stage, short-term rescheduling decisions are made:

- Calling on-call nurse
- Amount of overtime
- Amount of undertime

4.1.1 Two Stage Stochastic Programming Model

Parameters:

t: Total number of nurses in the department (assumed to be fixed during the month)

g: Total number of days to be scheduled

S: Total number of scenarios

$$m_i = \begin{cases} 1 & \text{if nurse i is a senior nurse} \\ 0 & \text{otherwise} \end{cases} \qquad \forall i = 1, .., t$$

$$w_{ij} = \begin{cases} 1 & \text{if nurse i has breast} - \text{feeding permission on day j} \\ 0 & \text{otherwise} \end{cases}$$

$$\forall i = 1,..,t \;;\; j = 1,..,g$$

d_{jk}^{s}: required number of nurses on day j at shift k according to scenario s.

$$\forall j = 1,..,g; \; k = 1,..,3$$

$$a_{ijk} = \begin{cases} 1 & \text{if nurse i has a request about not working on day j at shift k} \\ 0 & \text{otherwise} \end{cases}$$

$$\forall i = 1,..,t \;;\; j = 1,..,g; \; k = 1,..,3$$

First Stage Decision Variables:

$$x_{ijk} = \begin{cases} 1 & \text{if nurse i works on day j at shift k} \\ 0 & \text{otherwise} \end{cases}$$

$$\forall i = 1,..,t \;;\; j = 1,..,g; \; k = 1,..,3$$

$$z_{ijk} = \begin{cases} 1 & \text{if nurse i is an on} - \text{call nurse on day j at shift k} \\ 0 & \text{otherwise} \end{cases}$$

$$\forall i = 1,..,t \;;\; j = 1,..,g; \; k = 1,..,3$$

$$f_{ij} = \begin{cases} 1 & \text{if nurse i is on day} - \text{off on day j} \\ 0 & \text{otherwise} \end{cases} \qquad \forall i = 1,..,t \;;\; j = 1,..,g$$

y_i: the total number of normal working shifts for nurse i during the scheduling period $\forall i = 1,..,t$

n_i: the total number of night shifts for nurse i during the scheduling period

$$\forall i = 1,..,t$$

p_i: the total number ofassignments as an on $-$ call nurse for nurse i during the scheduling period $\forall i = 1,..,t$

Second Stage Decision Variables:

o_{jk}^{s}: required number of over time nurses on day j at shift k according to

scenario s $\quad \forall j = 1,..,g;\ k = 1,..,3;\ s = 1,..,S$

$$c_{jk}^{s} = \begin{cases} 1 & \text{if on} - \text{call nurse will be called on day j at shift k according to scenario s} \\ 0 & \text{otherwise} \end{cases}$$

$\forall j = 1,..,g;\ k = 1,..,3;\ s = 1,..,S$

r_{jk}^{s}: Number of nurses who will be permitted to go on day j at shift k according to

scenario s $\quad \forall j = 1,..,g;\ k = 1,..,3;\ s = 1,..,S$

Mathematical Formulation:

$$\mathbf{min}\ Q(x)$$

Subject to:

1. There must be at least one senior nurse in each shift.

$$\sum_{i=1}^{t} x_{ijk} m_i \geq 1 \qquad \forall j = 1,..,g\ ;\ k = 1,..,3 \qquad (1)$$

2. Nurses who are pregnant or have breast-feeding permission can not be assigned to any third shift (night shift).

$$w_{ij} + x_{ij3} + z_{ij3} \leq 1 \qquad \forall i = 1,..,t\ ;\ j = 1,..,g \qquad (2)$$

3. There must be at least two shift periods between sequential working shifts for each nurse.

$$x_{ij2} + x_{ij3} + x_{i(j+1)1} \leq 1 \qquad \forall i = 1,..,t\ ;\ j = 1,..,g \qquad (3.a)$$

$$x_{ij3} + x_{i(j+1)1} + x_{i(j+1)2} \leq 1 \qquad \forall i = 1,..,t\ ;\ j = 1,..,g-1 \qquad (3.b)$$

4. Exactly one on-call nurse must be assigned for each shift.

$$\sum_{i=1}^{t} z_{ijk} = 1 \qquad \forall j = 1,..,g\,;\, k = 1,..,3 \qquad (4)$$

5. Nurses who are on day-off can not be assigned as a normal working nurse. (They can not be assigned as an on-call nurse as well, this is guaranteed by constraint 9.)

$$\left(\sum_{k=1}^{3} x_{ijk}\right) + f_{ij} = 1 \qquad \forall i = 1,..,t\,;\, j = 1,..,g \qquad (5)$$

6. There must be two days-off in total in every week during scheduling period for each nurse.

$$\sum_{j=1}^{7} f_{ij} = 2 \qquad \forall i = 1,..,t \qquad (6.a)$$
$$\sum_{j=8}^{14} f_{ij} = 2 \qquad \forall i = 1,..,t \qquad (6.b)$$
$$\sum_{j=15}^{21} f_{ij} = 2 \qquad \forall i = 1,..,t \qquad (6.c)$$
$$\sum_{j=22}^{28} f_{ij} = 2 \qquad \forall i = 1,..,t \qquad (6.d)$$

7. Assigning nurses to 4 night shifts sequentially is not wanted.

$$x_{ij3} + x_{i(j+1)3} + x_{i(j+2)3} + x_{i(j+3)3} \leq 3 \quad \forall i, j = 1,..,t \; and \; j = 1,..,g-3 \qquad (7)$$

8. The difference between the number of night shifts, total regular shifts and number of assigments as an on-call nurse of nurses must be less than or equal to two.

$$y_i = \sum_{k=1}^{3} \sum_{j=1}^{g} x_{ijk} \qquad \forall i = 1,..,t \qquad (8.a)$$
$$n_i = \sum_{j=1}^{g} x_{ij3} \qquad \forall i = 1,..,t \qquad (8.b)$$

$p_i = \sum_{k=1}^{3} \sum_{j=1}^{g} z_{ijk}$ $\forall i = 1, .., t$ (8.c)

$p_i - p_j \leq 2$ $\forall i, j = 1, .., t$ and i≠j (8.d)

$n_i - n_j \leq 2$ $\forall i, j = 1, .., t$ and i≠j (8.e)

$y_i - y_j \leq 2$ $\forall i, j = 1, .., t$ and i≠j (8.f)

9.

- The on-call nurse for the first shift is chosen among the nurses who will work in the second shift on same day.
- The on-call nurse for the third shift is chosen among the nurses who have worked in the second shift on same day.
- The on-call nurse for the second shift is chosen among the nurses either who have worked in the first shift or will work in the third shift on same day.

$z_{ij1} \leq x_{ij2}$ $\forall i = 1, .., t\ ;\ j = 1, .., g$ (9.a)

$z_{ij3} \leq x_{ij2}$ $\forall i = 1, .., t\ ;\ j = 1, .., g$ (9.b)

$z_{ij2} \leq x_{ij1} + x_{ij3}$ $\forall i = 1, .., t\ ;\ j = 1, .., g$ (9.c)

10. If nurse i has a request about not working on day j at shift k, then nurse i can not be assigned that shift on that day.

$a_{ijk} + x_{ijk} + z_{ijk} \leq 1$ $\forall i = 1, .., t\ ;\ j = 1, .., g;\ k = 1, .., 3$ (10)

11. Sign constraints

$x_{ijk}, z_{ijk} \in \{0,1\}$ $\forall i = 1, .., t\ ;\ j = 1, .., g;\ k = 1, .., 3$ (11.a)

$f_{ij} \in \{0,1\}$ $\forall i = 1, .., t\ ; j = 1, .., g$ (11.b)

$p_i, n_i, y_i \geq 0$ $\forall i = 1, .., t$ (11.c) where $\boldsymbol{Q}(\boldsymbol{x}) = \boldsymbol{E}_{\boldsymbol{s}}\big(\boldsymbol{Q}(\boldsymbol{x},\boldsymbol{s})\big)$ and

$$\boldsymbol{Q}(\boldsymbol{x},\boldsymbol{s}) = \mathbf{min}\Big(\sum_{\boldsymbol{j=1}}^{\boldsymbol{g}}\sum_{\boldsymbol{k=1}}^{\boldsymbol{3}}(\boldsymbol{a} * \boldsymbol{o}_{\boldsymbol{jk}}^{\boldsymbol{s}} + \boldsymbol{b} * \boldsymbol{c}_{\boldsymbol{jk}}^{\boldsymbol{s}} + \boldsymbol{c} * \boldsymbol{r}_{\boldsymbol{jk}}^{\boldsymbol{s}})\Big)$$

Subject to:

12. There must be required number of nurses at each shift to satisfy demand.

$\sum_{i=1}^{t} x_{ijk} + o_{jk}^{s} + c_{jk}^{s} - r_{jk}^{s} = d_{jk}^{s}$ $\quad \forall j = 1,..,g; k = 1,..,3$ (12.a)

$c_{jk}^{s} \leq 1$ $\quad \forall j = 1,..,g;\ k = 1,..,3$ (12.b)

13. Sign constraints

$c_{jk}^{s}, o_{jk}^{s}, r_{jk}^{s} \geq 0$ $\quad \forall j = 1,..,g;\ k = 1,..,3$ (13)

Two settings are considered about the objective function coefficients in the model:

The first setting is that assignments of overtime and undertime nurses are more costly than the assignment of on-call nurse and there is no difference between overtime and undertime. As a result, the objective function coefficient of calling an on-call nurse (b) is lower than assigning overtime (a) or undertime (c) nurse ($a, b, c >= 0\ and\ a, c > b$). The cost of assigning overtime (a) and undertime (c) nurses are taken as 4 and the cost of calling on-call nurses (b) is taken as 2.

The second setting is that assignments of overtime and undertime nurses are more costly than the assignment of on-call nurses and the assignment of overtime nurses is more valuable than the assignment of undertime nurses. As a result, the objective function coefficient of calling an on-call nurse (b) is lower than assigning overtime (a) or undertime (c) nurse, and the objective function coefficient of assigning undertime nurse (c) is lower than assigning overtime nurse (a) ($a, b, c >= 0\ and\ a > c > b$). The cost of assigning overtime nurses (a) is taken as 6, the cost of assigning undertime nurses (c) is taken as 4 and the cost of calling on-call nurses (b) is taken as 2.

4.1.2 L-Shaped Method

Since the second-stage problem is feasible under every feasible first-stage solution, we only use optimality cuts in our L-shaped method. Due to the nice structure of our second-stage problem, it is possible to obtain the dual variables without solving a linear program. In our computational study, we use both methods (i.e. solving the second-stage model as an LP and using the closed form solutions of the dual variables) to illustrate the improvement brought by using the closed form solutions.

Let y_{jk}, z_{jk} be the dual variables of optimality subproblem. Then dual problem is separable across j and k, and the closed form of dual problem can be obtained as follows:

$$\max \hat{d}_{jk} * y_{jk} + z_{jk}$$

Subject to:

$$-c \le y_{jk} \le a$$

$$y_{jk} + z_{jk} \le b$$

$$y_{jk} \; urs; \; z_{jk} \le 0$$

where $\hat{d}_{jk} = d_{jk} - \sum_{i=1}^{t} \hat{x}_{ijk}$ where $\hat{x}_{ijk}$ is the first-stage solution in the considered iteration. The feasible region of the dual problem is given in Figure 4.1. Accordingly, the optimal solution of the dual problem (y_{jk}, z_{jk}) is:

- If $\hat{d}_{jk} \ge 1$,then the optimal solution will be $(y_{jk} = a, z_{jk} = b - a)$,
- If $0 < \hat{d}_{jk} < 1$,then the optimal solution will be $(y_{jk} = b, z_{jk} = 0)$,
- If $\hat{d}_{jk} \le 0$, then the optimal solution will be $(y_{jk} = -c, z_{jk} = 0)$.

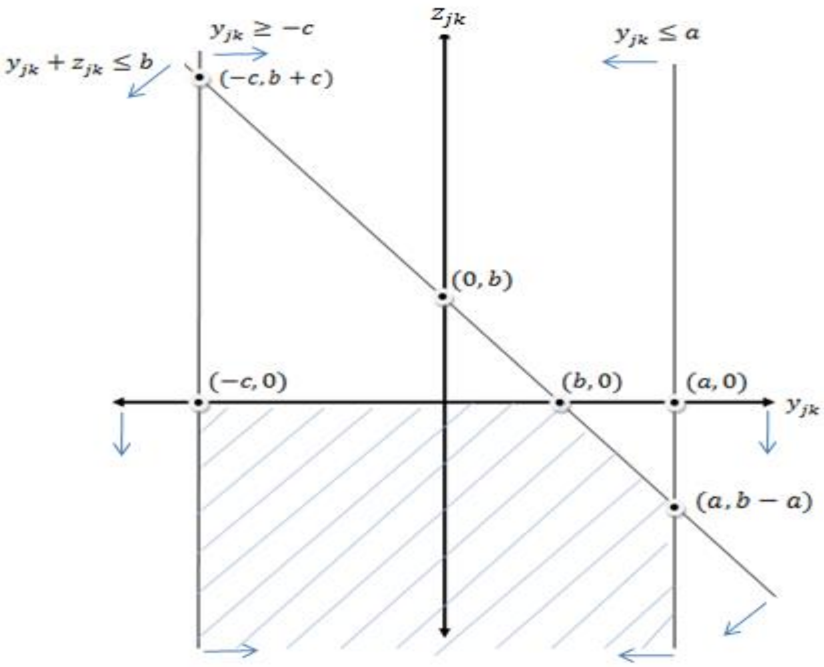


Figure 4.1 The Feasible Region of The Dual Problem

In addition, a lower bound is added to the master problem by using mean value solution since it is expected to be:

$$EV \leq RP$$

To add this bound new variables are defined:

v_{jk}: *required number of over time nurses on day j at shift k according to average scenario.* $\forall j = 1, .., g;\ k = 1, .., 3$

u_{jk}: *required number of under time nurses on day j at shift k according to average scenario.* $\forall j = 1, .., g;\ k = 1, .., 3$

$$n_{jk} = \begin{cases} 1 & if\ on-call\ nurse\ will\ be\ called\ on\ day\ j\ at\ shift\ k\ according\ to \\ & average\ scenario \\ 0 & otherwise \end{cases}$$

$$\forall j = 1,..,g;\ k = 1,..,3$$

And the constraints given below is added to the master problem.

$$\theta \geq \sum_{j=1}^{g} \sum_{k=1}^{3} a * v_{jk} + b * n_{jk} + c * u_{jk} \quad (14)$$

where θ approximates the expected cost of the second stage.

$$\sum_{i=1}^{t} x_{ijk} + v_{jk} + n_{jk} - u_{jk} = \bar{d}_{jk} \qquad \forall j = 1,..,g;\ k = 1,..,3 \quad (15)$$

where $\bar{d}_{jk}$ represents the average required number of nurses. Constraint (14) represents that $RP \geq EV$. Constraint (15) ensures meeting the demand in mean value problem.

In addition to single-cut approach, multi-cut approach is used as an alternative solution approach while adding optimality cuts. While in single-cut approach only one cut for all realizations is added at each iteration, in multi-cut approach, one cut for each realization is added at each iteration. As a result, master problem becomes larger at each iteration according to the single-cut approach. Birge and Louveaux [4] mention that the multi-cut approach is expected to be more efficient than single-cut approach when the number of realizations is not significantly larger than the number of first-stage constraints.

In classical L-shaped algorithm, the master problem is solved, then feasibility and optimality subproblems are solved and feasibility and optimality cuts associated with the current solution are added to the master problem and master problem is solved again. This process repeats iteratively. At every solution of master problem a new search tree is constructed from the beginning. On the other hand, L-shaped based branch-and-cut approach applies the algorithm on a single search tree. This approach is applied by using the lazy constraint callback feature of CPLEX. Feasibility and optimality subproblems are solved for each integer feasible node in the branch-and-bound tree of the master problem.

4.2 Forecasting Demand

Historical data of 9 months of daily TISS is used. We use time series analysis to forecast the demand. The analysis is made in EViews 8. After selection the best model, the confidence intervals related to each day of forecasting period are found and scenarios generations are made according to these confidence intervals.

4.2.1 The Run Sequence Plot

The run- sequence plot for the daily nurse requirement during 9 months is given in Figure 4.2. Original series is called as "nurse_count".

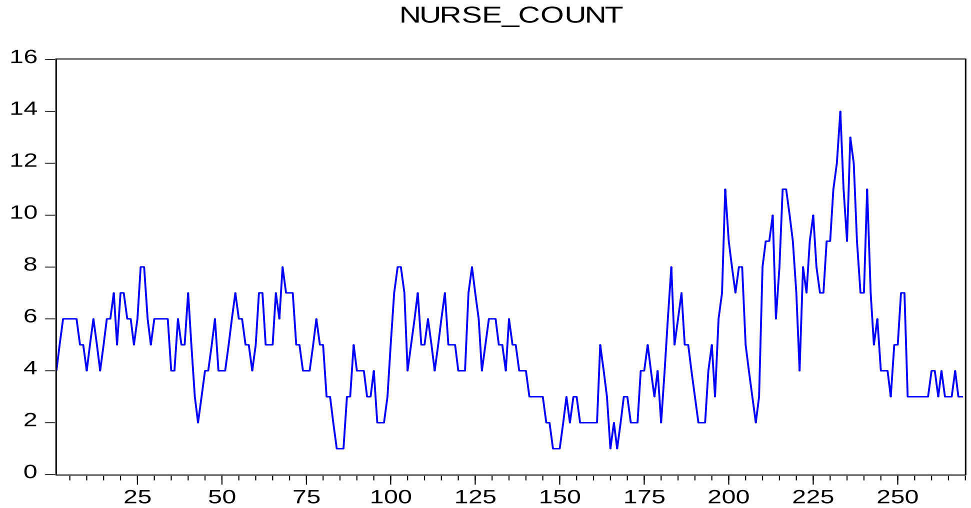


Figure 4.2 The Required Number of Nurses During 9 Months (Original Series)

4.2.2 Unit Root Test

Augmented Dickey-Fuller Unit Root Test is used to test the stationarity of data. Test results are given in Table 4.1. More details about results can be found in Appendix B.

Table 4.1 Unit Root Test Results

Test for Unit Root in	**Include in Test Equation**	**t- statistics**	**Probability**
Level	Intercept	-5.042117	0.0000
Level	Intercept and Trend	-5.032603	0.0002

In order to test stationarity, hypothesized as:

$H_0 = Data\ is\ not\ stationary$

$H_1 = Data\ is\ stationary$

Since probabilities are less than the α value, which is 0.05 for the 95% confidence interval, H_0 is rejected. In other words, data is said to be stationary.

4.2.3 ACF and PACF (Correlogram)

As explained in review of methodologies, ACF and PACF give a prior knowledge about the model for stationary data. Correlogram for "nurse_count" series is given in Figure 4.3. It can be seen that ACF gradually decreases, and PACF cuts off after 1 lag. Initial interpretation is in the direction of usage of AR(1) model.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.823	0.823	184.78	0.000
		2	0.669	-0.023	307.52	0.000
		3	0.567	0.072	396.07	0.000
		4	0.505	0.074	466.56	0.000
		5	0.476	0.093	529.39	0.000
		6	0.442	0.010	583.82	0.000
		7	0.431	0.094	635.71	0.000
		8	0.391	-0.058	678.50	0.000
		9	0.346	-0.004	712.24	0.000
		10	0.308	-0.004	739.01	0.000
		11	0.326	0.160	769.07	0.000
		12	0.360	0.077	805.95	0.000
		13	0.369	0.013	844.87	0.000
		14	0.364	0.013	882.94	0.000
		15	0.314	-0.097	911.32	0.000
		16	0.278	0.014	933.69	0.000
		17	0.220	-0.107	947.73	0.000
		18	0.184	-0.005	957.56	0.000
		19	0.167	-0.018	965.67	0.000
		20	0.142	-0.033	971.62	0.000
		21	0.129	0.017	976.49	0.000
		22	0.082	-0.080	978.46	0.000
		23	0.043	-0.037	979.00	0.000
		24	0.010	-0.037	979.03	0.000
		25	-0.018	-0.068	979.13	0.000
		26	-0.057	-0.100	980.09	0.000
		27	-0.056	0.072	981.03	0.000
		28	-0.063	-0.058	982.23	0.000
		29	-0.066	0.058	983.56	0.000
		30	-0.066	0.005	984.87	0.000
		31	-0.071	0.031	986.41	0.000
		32	-0.084	-0.062	988.58	0.000
		33	-0.061	0.138	989.74	0.000
		34	-0.033	0.032	990.08	0.000
		35	-0.055	-0.118	991.02	0.000
		36	-0.065	0.057	992.36	0.000

Figure 4.3 ACF and PACF of the "nurse_count" Series

4.2.4 Selection of Best Model

4.2.4.1 Models

In order to select the best appropriate model, the forecasted model should have the following features at the same time:

i. Model must be statistically significant.

ii. The forecasted coefficients of model must be statistically significant.

iii. The forecasted coefficients of the model must be in the range.

iv. The model must have the least AIC and SIC values.

The related information about the comparison of models are given in Table 4.2.

Table 4.2 Comparison of Models

Model	**Significance of Model**	**Coefficient of Determination (R^2)**	**Variable**	**Coefficient**	**Probab**	**Information Criterion**
AR(1)	F=565.5900 p=0.00	0.679314	AR(1)	0.825074	0.000	AIC=3.441838
						SIC=3.46856
AR(2)	F=281.4766 p=0.000000	0.679934	AR(1)	0.845326	0.000	AIC=3.451141
			AR(2)	-0.023894	0.6977	SIC=3.49133
MA(1)	F=264,0596 p=0.000000	0.496297	MA(1)	0.708676	0.000	AIC=3.890378
						SIC=3.917033
MA(2)	F=198.5845 p=0.000000	0.597994	MA(1)	0.822148	0.0000	AIC=3.672266
			MA(2)	0.440838	0.0000	SIC=3.712249
ARMA (1,1)	F=282.0376 p=0.000000	0.679547	AR(1)	0.813562	0.0000	AIC=3.448546
			MA(1)	0.035745	0.6309	SIC=3.488635
ARMA (2,1)	F=189.1090 p=0.000000	0.682435	AR(1)	-0.043343	0.7480	AIC=3.450757
			AR(2)	0.705077	0.0000	
			MA1)	0.908738	0.0000	SIC=3.504354
ARMA (1,2)	F=189.1844 P=0.000000	0.681702	AR(1)	0.871690	0.0000	AIC=3.449233
			MA(1)	-0.041987	0.5889	
			MA(2)	-0.128356	0.0822	SIC=3.502686
ARMA (2,2)	F=141.3767 p=0.000000	0.682561	AR(1)	0.007282	0.9728	AIC=3.457822
			AR(2)	0.678531	0.0000	
			MA(1)	0.844894	0.0002	SIC=3.524818
			MA(2)	-0.031122	0.7234	

According to Table 4.2 it is seen that all models are statistically significant. If we evaluate each model separately:

The forecasted coefficient of AR(1) model, ϕ_1, is statistically significant and between the range of [-1,1]. As a result AR(1) model is said to be appropriate.

The forecasted coefficients of AR(2) model are ϕ_1 and ϕ_2. While the coefficient of ϕ_1 is statistically significant, ϕ_2 is not significant. So AR(2) model is not an appropriate model.

The forecasted coefficient of MA(1) model, which is θ_1, is statistically significant and between the range of [-1, 1]. Therefore MA(1) model is an appropriate model.

The forecasted coefficients of MA(2) model, which are θ_1 and $\theta_2$are statistically significant and between the range of [-1, 1]. For this reason, MA(2) model is an appropriate model.

The forecasted coefficients of ARMA(1,1) model, ϕ_1, is statistically significant, on the other hand the other forecasted coefficient of model, θ_1, is not significant. Therefore ARMA(1,1) model is not an appropriate model.

ARMA(2,1) is not an appropriate model because while the coefficients ϕ_2 and θ_1 are statistically significant, ϕ_1 is not statistically significant.

ARMA(1,2) is not an appropriate model because while the forecasted coefficient ϕ_1 is statistically significant, the forecasted coefficients θ_1 and θ_2 are not statistically significant.

ARMA(2,2) is not an appropriate model because while the coefficients ϕ_2 and θ_1 are statistically significant, ϕ_1 and θ_2 are not statistically significant.

The more details about forecasted models can be found in Appendix C.

AR(1), MA(1) and MA(2) models satisfy the first three features as a result the model with the smallest information criterion values will be chosen as the best appropriate model. Since the AR(1) has the smallest AIC and SIC values, AR(1) is chosen as the forecasting model.

4.2.4.2 Ljung-Box Statistics

In order to test the efficiency of the AR(1) model, Ljung-Box Statistics is used by examination of correlogram for residuals. The ACF and PACF related to residuals are given in Figure 4.4.

Date: 09/13/15 Time: 13:56
Sample: 1 270
Included observations: 269

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.020	0.020	0.1116	0.738
		2	-0.065	-0.066	1.2824	0.527
		3	-0.043	-0.041	1.8008	0.615
		4	-0.039	-0.041	2.2106	0.697
		5	0.059	0.055	3.1726	0.673
		6	-0.014	-0.023	3.2267	0.780
		7	0.118	0.125	7.1265	0.416
		8	0.045	0.042	7.7016	0.463
		9	0.020	0.039	7.8106	0.553
		10	-0.113	-0.107	11.398	0.327
		11	-0.011	0.012	11.436	0.408
		12	0.099	0.075	14.194	0.289
		13	0.070	0.064	15.596	0.272
		14	0.155	0.146	22.480	0.069
		15	-0.004	0.012	22.485	0.096
		16	0.086	0.112	24.620	0.077
		17	-0.031	-0.009	24.905	0.097
		18	-0.023	0.010	25.056	0.123
		19	0.038	0.012	25.469	0.146
		20	-0.020	-0.043	25.582	0.180
		21	0.096	0.049	28.295	0.132
		22	-0.018	-0.016	28.388	0.163
		23	-0.011	-0.014	28.422	0.200
		24	-0.006	0.008	28.434	0.242
		25	0.029	0.025	28.686	0.277
		26	-0.098	-0.138	31.591	0.207
		27	0.009	-0.001	31.614	0.247
		28	-0.012	-0.098	31.660	0.289
		29	-0.003	-0.015	31.662	0.335
		30	-0.007	-0.069	31.678	0.383
		31	0.013	0.044	31.728	0.430
		32	-0.095	-0.143	34.513	0.349
		33	-0.013	0.005	34.566	0.393
		34	0.121	0.120	39.099	0.251
		35	-0.052	-0.074	39.953	0.259
		36	0.016	0.028	40.033	0.296

Figure 4.4 ACF and PACF Related to Residuals

The hypothesis of Ljung-Box statistics for the AR(1) model is:

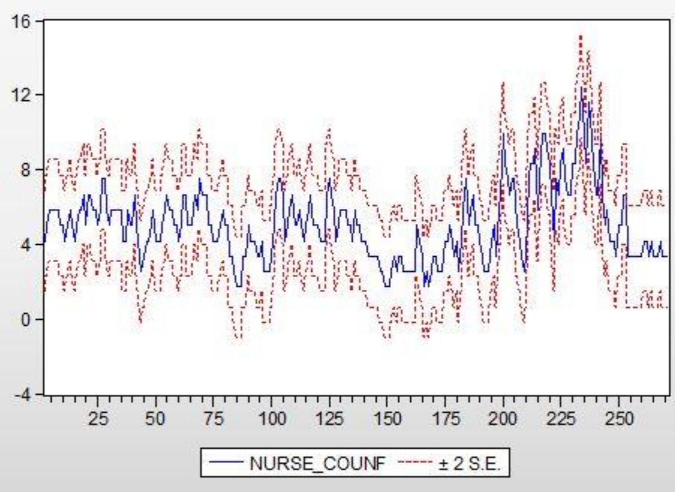
$H_0 = Residuals\ are\ not\ linearly\ dependent, they\ show\ random\ distribution$

$H_1 = Residuals\ are\ linearly\ dependent, they\ do\ not\ show\ random\ distribution$

Since $p = 0.296 > \alpha = 0.05$, the hypothesis H_0 is accepted. It is concluded that residuals are not linearly dependent and they show random distribution and for this reason AR(1) model is an appropriate and efficient model for forecasting.

4.2.5 Forecasting

In order to test the validity of model, prediction is made and the results given below in Figure 4.5 and Figure 4.6 are obtained.



Forecast: NURSE_COUNF	
Actual: NURSE_COUNT	
Forecast sample: 1 271	
Adjusted sample: 2 271	
Included observations: 269	
Root Mean Squared Error	1.342516
Mean Absolute Error	1.027548
Mean Abs. Percent Error	24.85948
Theil Inequality Coefficient	0.121322
Bias Proportion	0.000000
Variance Proportion	0.096368
Covariance Proportion	0.903632

Figure 4.5 Prediction Result of the Forecasted Model

Since Theil's Inequality Coefficient and the Mean Absolute Percentage Error show the success of the forecasted model, these two measurements are examined. The Theil Inequality Coefficient is expected to be in the range of [0,1] and being close tozero is desirable and MAPE is wanted to take a small value for the success of the model. Since Theil's Inequality Coefficient is 0.121322 which is close to zero and MAPE is 24.85948%, AR(1) model is found successful.

In Figure 4.6, the predictions made by forecasted model and the actual values for requirement of nurses during 9 months are shown. It is seen that there is a high consistency which supports the usage of AR(1) model.

The prediction made by AR(1) model for the next period is 3.36 nurses and the standard deviation is $\pm 2\ S.E.$ where $S.E.$ is equal to 1.34, which can be seen in Figure C.1, as a result the required number of nurses for the next period is expected to be in the range of 3.36 ± 2.68, in other words [0.68, 6.04]. Since the required number of nurses is always be integer and satisfying the all demand is essential, the estimated required number of nurses for the next period is in the range of [1, 7].

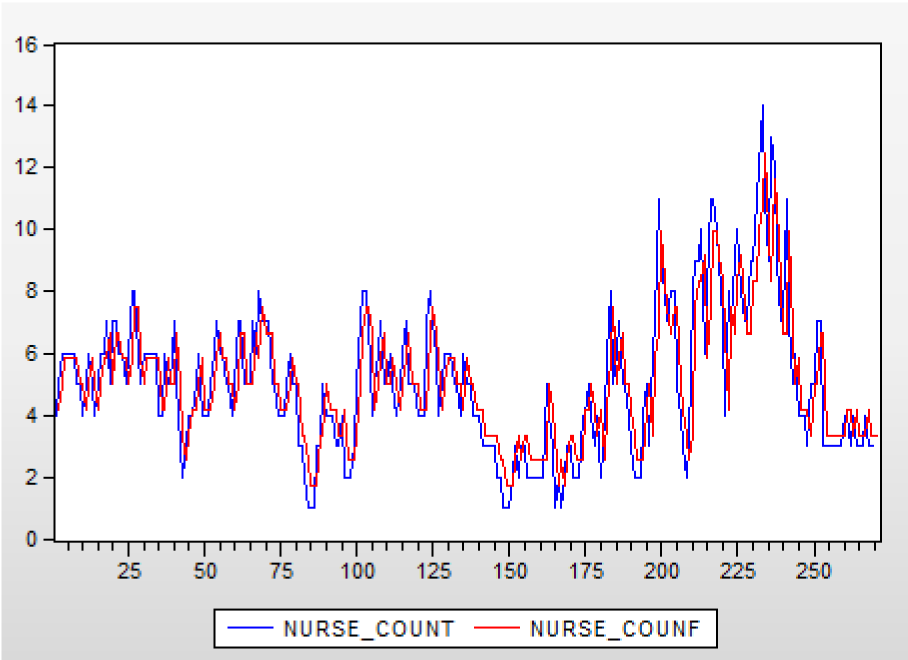


Figure 4.6 Predictions and Actual Values

CHAPTER 5

COMPUTATIONAL RESULTS AND COMPARISON

By using our numerical results we aim to analyze the efficiency of our solution methods and the value of considering uncertainty in the problem. Objective function values, solution times of each instance and the uncertainty related measures (which are the Expected Value of Perfect Information and the Value of the Stochastic Solution) are presented in the following parts.

First step is testing the performance of solution methods by measuring solution times. In order to do this, we generate scenarios with different scheduling periods. Our data structure, which can be seen in Figure 5.1 includes 9 months and 36 weeks. In Figure 5.1, "M" represents months, "TW" represents two-week long periods and "W" represents weeks. For each day, required number of nurses is available as historical data and the number of nurses working in the unit is also known. The number of senior nurses and the number of special requests and breast-feeding permissions are assigned according to the interviews made by the head nurse.

M1				M2				M3				M4				M5				M6				M7				M8				M9			
TW1		TW2		TW3		TW4		TW5		TW6		TW7		TW8		TW9		TW10		TW11		TW12		TW13		TW14		TW15		TW16		TW17		TW18	
W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4

Figure 5.1 Data Structure

8 instances, which include scheduling periods of 4, 5, 6, 7 and 14 days are generated initially. The first 4, 5, 6 and 7 days of the 33rd week (first week of 9th month which

is shown in Figure 5.2) are taken as forecasting periods randomly in order to generate scenarios for 4, 5, 6 and 7 days long scheduling periods.

M1				M2				M3				M4				M5				M6				M7				M8				M9			
TW1		TW2		TW3		TW4		TW5		TW6		TW7		TW8		TW9		TW10		TW11		TW12		TW13		TW14		TW15		TW16		TW17		TW18	
W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	**W1**	W2	W3	W4

Figure 5.2 Forecasting Period for the 4, 5, 6 and 7 Days Long Schedules

The 95% confidence intervals for these periods are given in Table 5.1. The scenarios are generated in the confidence intervals uniformly.

Table 5.1 95% Confidence Intervals for the 4, 5, 6 and 7 Days Long Scheduling Periods

		Scheduling Period			
		4 Days	5 Days	6 Days	7 Days
Day	1	[6,12]	[6,12]	[6,12]	[6,12]
	2	[7,12]	[7,12]	[7,12]	[7,12]
	3	[5,11]	[5,11]	[5,11]	[5,11]
	4	[4,10]	[4,10]	[4,10]	[4,10]
	5		[4,10]	[4,10]	[4,10]
	6			[6,12]	[6,12]
	7				[6,12]

In the same way, in order to generate scenarios for the 14 days long schedule, the first 14 days of the 9th month, which is shown in Figure 5.3, is taken randomly as forecasting period.

M1				M2				M3				M4				M5				M6				M7				M8				M9			
TW1		TW2		TW3		TW4		TW5		TW6		TW7		TW8		TW9		TW10		TW11		TW12		TW13		TW14		TW15		TW16		**TW17**		TW18	
W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4

Figure 5.3 Forecasting Period for the 14 Days Long Schedule

In the tables below *g* represents the number of days to be scheduled, *t* represents the number of nurses to be scheduled and *s* represents the number of scenarios. Time is given in terms of CPU seconds and "-" represents that any solution could not be obtained in 3 hours and "*" represents the best solution obtained at the end of 3 hours.

Our computational experiments are executed on a 3.10 GHz computer with 16.0 GB memory and 64-bit Windows 7 operating system. CPLEX 12.6.1 is used as the solver. Solution times of each method for the 4, 5, 6, 7 and 14 days long schedules are given in Table 5.2.

Table 5.2 Solution Times of Each Approach for Each Instance

Instance	***g/t/s***	**Extensive**	**L-shaped**	**L-Shaped with Dual Solution**	**L-Shaped with dual solution and lower bound**	**Multi-Cut L-shaped with dual solution and lower bound**	**L-shaped based branch and cut**
1	4/4/2	0.23	0.43	0.27	0.32	0.35	0.04
2	4/4/10	0.13	0.45	0.29	0.27	0.33	0.05
3	5/5/2	0.22	0.31	0.27	0.22	0.57	0.12
4	5/5/10	0.31	0.41	0.40	0.38	0.48	0.12
5	6/6/2	0.24	0.51	0.39	0.28	0.53	0.13
6	6/6/10	0.31	0.7	0.6	0.47	-	0.15
7	6/7/10	0.4	0.66	0.43	0.8	0.91	0.18
8	7/5/10	0.23	0.31	0.23	0.23	-	0.1

According to Table 5.2, the following inferences are made:

- For all instances, solving subproblems by closed form of dual solution improves the solution performance in terms of time.

- Adding lower bound improves the solution performance in terms of time for instances 2, 3, 4 and 5.
- Single-cut approach outperforms the multi-cut approach in all instances and it is observed that while the optimal solutions for instance 6 and instance 8 can not be obtained by multi-cut approach, single-cut approach provides the optimal solution for these instances.
- Extensive form and L-shaped based branch-and-cut methods outperform all other solution approaches in all instances.

To determine the method will be used to perform the main runs of the numerical experiment, the best two approaches, which are extensive form and L-shaped based branch and cut, are compared. In order to test the performance of extensive form and L-shaped based branch and cut, eighteen different instances are generated. These instances and the solution times of these methods for each instance are given in Table 5.3.

Table 5.3 Solution Times of Extensive Form and L-shaped Based Branch and Cut

Instance	***g/t/s***	**Solution Time of Extensive Form**	**Solution Time of L-shaped Based Branch and Cut**
9	7/5/10	0.23	0.10
10	7/5/50	0.15	0.10
11	7/5/100	0.40	0.11
12	7/5/200	0.32	0.11
13	7/6/10	0.10	0.31
14	7/7/10	0.30	0.13
15	14/8/10	0.25	0.25
16	14/9/10	0.34	0.27
17	14/10/10	0.54	0.28
18	1411/10	0.53	0.48
19	14/12/10	0.60	0.33
20	14/13/10	0.70	4.20
21	14/14/10	8.40	8.50
22	14/15/10	0.73	12.94
23	14/16/10	0.90	157.83
24	14/17/10	0.88	342.57
25	28/17/100	130	28654.98
26	28/17/200	8.68	-

According to Table 5.3, the following inferences are made:

- For instance 9, L-shaped based branch and cut performs better than the extensive form.
- As number of scenarios is increased, L-shaped based branch and cut still gives optimal solution in a shorter time in comparison with extensive form.
- When the number of days is increased to 14 days and the number of nurses is increased to 8, L-shaped based branch and cut still gives optimal solution in a shorter time in comparison with extensive form.
- When the number of nurses is increased for the 14 days long schedule, it is observed that until instance 20 with 13 nurses, L-shaped based branch-and-

cut outperforms the extensive form. On the other hand, beginning from the instance 20, it is observed that the extensive form starts to outperform L-shaped based branch-and-cut method. In addition to this, if the problem size gets bigger, the gap between the performances of two methods increases.

Based on the results of our preliminary runs, we use the extensive form to perform the main runs of our numerical experiment.

In order to test the performance of extensive form at the actual size of problem, each month is taken as the forecasting period separately as seen in Figure 5.4.

<table>
<tr><td colspan="4">M1</td><td colspan="4">M2</td><td colspan="4">M3</td><td colspan="4">M4</td><td colspan="4">M5</td><td colspan="4">M6</td><td colspan="4">M7</td><td colspan="4">M8</td><td colspan="4">M9</td></tr>
<tr><td colspan="2">TW1</td><td colspan="2">TW2</td><td colspan="2">TW3</td><td colspan="2">TW4</td><td colspan="2">TW5</td><td colspan="2">TW6</td><td colspan="2">TW7</td><td colspan="2">TW8</td><td colspan="2">TW9</td><td colspan="2">TW10</td><td colspan="2">TW11</td><td colspan="2">TW12</td><td colspan="2">TW13</td><td colspan="2">TW14</td><td colspan="2">TW15</td><td colspan="2">TW16</td><td colspan="2">TW17</td><td colspan="2">TW18</td></tr>
<tr><td>W1</td><td>W2</td><td>W3</td><td>W4</td><td>W1</td><td>W2</td><td>W3</td><td>W4</td><td>W1</td><td>W2</td><td>W3</td><td>W4</td><td>W1</td><td>W2</td><td>W3</td><td>W4</td><td>W1</td><td>W2</td><td>W3</td><td>W4</td><td>W1</td><td>W2</td><td>W3</td><td>W4</td><td>W1</td><td>W2</td><td>W3</td><td>W4</td><td>W1</td><td>W2</td><td>W3</td><td>W4</td><td>W1</td><td>W2</td><td>W3</td><td>W4</td></tr>
</table>

Figure 5.4 Forecasting Periods for the Four Weeks Long Schedules

Forecasting for each month is made and confidence intervals for the days of each month are found and scenarios are generated according to these confidence intervals. The confidence intervals related to each month are given in Table5.4.

Table 5.4 95% Confidence Intervals for the Four Weeks Long Scheduling Periods

		Scheduling Period								
		1st Month	2nd Month	3rd Month	4th Month	5th Month	6th Month	7th Month	8th Month	9th Month
Day	1	[2,7]	[4,9]	[4,9]	[1,5]	[3,8]	[2,7]	[1,6]	[1,7]	[6,12]
	2	[3,8]	[3,8]	[3,8]	[1,5]	[2,7]	[1,7]	[1,7]	[4,9]	[7,12]
	3	[4,9]	[4,9]	[3,8]	[1,5]	[3,8]	[1,7]	[1,7]	[4,10]	[5,11]
	4	[4,9]	[4,9]	[2,7]	[1,7]	[4,9]	[1,7]	[1,6]	[8,13]	[4,10]
	5	[4,9]	[4,9]	[3,8]	[1,7]	[4,10]	[1,7]	[1,6]	[6,12]	[4,10]
	6	[4,9]	[4,9]	[4,10]	[3,8]	[3,8]	[1,7]	[1,6]	[5,11]	[6,12]
	7	[4,9]	[4,9]	[4,10]	[2,7]	[3,8]	[1,6]	[2,7]	[4,10]	[6,12]
	8	[3,8]	[2,7]	[3,8]	[2,7]	[3,8]	[1,6]	[2,7]	[5,11]	[8,13]
	9	[3,8]	[2,7]	[3,8]	[2,7]	[2,7]	[1,5]	[3,8]	[5,11]	[9,14]
	10	[2,7]	[4,9]	[3,8]	[1,7]	[2,7]	[1,5]	[2,7]	[3,8]	[10,16]
	11	[3,8]	[3,8]	[4,10]	[1,7]	[2,7]	[1,5]	[1,7]	[2,7]	[8,13]
	12	[4,9]	[3,8]	[4,9]	[2,7]	[4,10]	[4,10]	[2,7]	[1,7]	[6,12]
	13	[3,8]	[4,10]	[5,11]	[1,6]	[5,11]	[5,11]	[1,6]	[1,6]	[9,15]
	14	[2,7]	[3,8]	[4,10]	[1,6]	[4,10]	[4,10]	[2,7]	[1,7]	[9,14]
	15	[3,8]	[1,7]	[4,10]	[1,6]	[4,9]	[4,9]	[4,9]	[5,11]	[6,12]
	16	[4,9]	[1,6]	[4,10]	[1,7]	[2,7]	[2,7]	[5,11]	[6,12]	[4,10]
	17	[4,9]	[1,7]	[3,8]	[3,8]	[3,8]	[3,8]	[3,8]	[6,12]	[4,10]
	18	[4,10]	[2,7]	[3,8]	[4,10]	[4,9]	[4,9]	[4,9]	[7,12]	[8,13]
	19	[3,8]	[2,7]	[2,7]	[5,11]	[4,9]	[4,9]	[4,10]	[4,9]	[4,10]
	20	[4,10]	[3,8]	[2,7]	[5,11]	[4,9]	[4,9]	[3,8]	[5,11]	[3,8]
	21	[4,10]	[4,9]	[2,7]	[4,10]	[3,8]	[3,8]	[3,8]	[8,13]	[4,9]
	22	[4,9]	[2,7]	[3,8]	[2,7]	[3,8]	[3,8]	[2,7]	[8,13]	[2,7]
	23	[4,9]	[2,7]	[4,9]	[3,8]	[2,7]	[2,7]	[1,7]	[7,12]	[2,7]
	24	[3,8]	[2,7]	[3,8]	[4,9]	[4,9]	[4,9]	[1,6]	[6,12]	[2,7]
	25	[4,9]	[3,8]	[3,8]	[4,10]	[3,8]	[3,8]	[1,6]	[4,10]	[1,7]
	26	[5,11]	[4,9]	[1,7]	[3,8]	[3,8]	[3,8]	[1,6]	[2,7]	[3,8]
	27	[5,11]	[4,10]	[1,7]	[3,8]	[2,7]	[2,7]	[2,7]	[5,11]	[3,8]
	28	[4,9]	[4,9]	[1,6]	[4,9]	[2,7]	[2,7]	[3,8]	[4,10]	[4,10]

As a result, 9 different instances with 100 scenarios are generated according to the forecasts of each month and each instance is represented as 27.1, 27.2, 27.3, etc., in

Table 5.5. The required number of nurses during a month on average for each instance is given with the last column of Table 5.5**.** In addition to these instances, which emphasize randomness, 9 different instances with 3 scenarios are generated. The first scenarios include the minimum values of confidence intervals, the second scenarios include the maximum values of confidence intervals and the third scenarios include the average value of confidence interval limits. These instances, which are given in Table 5.5 as instance 28.1, 28.2, etc., are generated to strongly reflect the time series connection between two consecutive days.

Table 5.5 Problem Instances with Different Number of Scenarios for Four Week Schedule

Instance	***g/t/s***	**Number of senior nurses**	**Number of breast-feeding permissions (number of days)**	**Number of special requests (number of shifts)**	**Total demand during a month on average (number of nurses)**
27.1	28/17/100	17	3	5	520.41
27.2	28/17/100	17	3	5	468.39
27.3	28/17/100	17	3	5	481.53
27.4	28/17/100	17	3	5	418.92
27.5	28/17/100	17	3	5	475.65
27.6	28/17/100	17	3	5	315.96
27.7	28/17/100	17	3	5	395.07
27.8	28/17/100	17	3	5	618.45
27.9	28/17/100	17	3	5	674.28
28.1	28/17/3	17	3	5	517.50
28.2	28/17/3	17	3	5	465
28.3	28/17/3	17	3	5	478.50
28.4	28/17/3	17	3	5	418.50
28.5	28/17/3	17	3	5	477
28.6	28/17/3	17	3	5	312
28.7	28/17/3	17	3	5	393
28.8	28/17/3	17	3	5	616.50
28.9	28/17/3	17	3	5	673.50

Table 5.6 and Table 5.7 shows the solution times for extensive form and uncertainty related measures for four week schedules. The last two rows include average values and maximum values of each measure over 9 instances.

Table 5.6 Solution Times, EVPI and VSS Measures for the Instances with 100 Scenarios

Instance	**Solution Time for Extensive Form**	**Objective Value of Stochastic Solution**	**Over time**	**Under time**	**On-call**	**EEV**	**WS**	**VSS (%)**	**EVPI (%)**
27.1	4.5	600.2	110.4	9.1	61.1	642.1	496.4*	6.5	17.3
27.2	131.5	528.6	81.5	24.1	53.0	544.2	346.2*	2.9	34.5
27.3	4538.9	562.5	93.7	19.0	55.9	585.3*	410.5*	3.9	27.0
27.4	47.7	494.2	58.7	42.5	44.7	506.3	235.5*	2.4	52.4
27.5	679.1	523.3	84.7	18.8	54.8	535.8	368.4*	2.3	29.6
27.6	4.9	444.7	23.2	71.7	32.5	446.2	45.2*	0.3	89.8
27.7	4.5	508.6	57.4	47.6	44.3	524.3	208.4*	3.0	59.0
27.8	507.8	957.7	195.8	7.9	71.6	1017	886.2*	5.8	7.5
27.9	4.8	1239.0	260.7	13.2	71.8	1276	1153.0	3.0	7.0
Avg.	658.2	651.0	107.3	28.2	54.4	675.3	461.1	3.4	36.0
Max	4538.9	1239.0	260.7	71.7	71.8	1276	1153.0	6.5	89.8

According to Table 5.6 and Table 5.7 the following inferences are made:

- The values of the stochastic solutions are in the range of 0.33%-6.53% and the expected values of the perfect information are in the range of 7%-89.8%. The value of stochastic solution is 3.35% on average and the expected value of perfect information is 36.01% on average.

- The expected value of perfect information takes the largest value and the value of the stochastic solution takes the smallest value for instance 19.6. From Table 5.5, it can be seen that instance 19.6 has the smallest demand on average during the month.

Table 5.7 Solution Times, EVPI and VSS Measures for the Instances with 3 Scenarios

Instance	Solution Time for Extensive Form	Objective Value of Stochastic Solution	Over time	Under time	On-call	EEV	WS	VSS (%)	EVPI (%)
28.1	2.3	683.0	124.0	19.3	54.8	740.3	596.0	7.7	12.7
28.2	7200.0	618.8*	94.2	35.7	50.5	636.3	465.3	2.8	24.8
28.3	2534.5	664.9	109.0	31.3	51.8	696.4*	528.7	4.5	20.5
28.4	2.8	606.7	74.8	56.0	41.7	622.7	373.3	2.6	38.5
28.5	2.6	637.0	102.5	30.3	52.8	654.3	504.7	2.7	20.8
28.6	2.6	558.7	37.3	87.3	30.0	562.7	177.7	0.7	68.2
28.7	3.1	602.0	72.5	57.8	40.3	620.7	333.7	3.0	44.6
28.8	4.4	1003.3	204.2	13.3	66.7	1113.3	940.0	9.9	6.3
28.9	2.0	1269.7	267.3	15.3	69.5	1357.0	1188.0	6.4	6.4
Avg.	85.6	738.2	124.0	38.9	51.0	778.2	567.5	4.5	27.3
Max	664.9	2534.5	267.3	87.3	69.5	1357.0	1188.0	9.9	68.2

According to Table 5.6 and Table 5.7 the following inferences are made:

- The average of the expected values of perfect information is higher in the instances with 100 scenarios than instances with 3 scenarios, therefore we can conclude that randomness of data has a significant impact on the value of the perfect information.

- The average of the values of the stochastic solutions is higher in the instances with 3 scenarios than the instances with 100 scenarios. Therefore, we can conclude that time series structure of the uncertainty has an impact on the value of the stochastic solution.
- The expected values of perfect information of the last two instances are the minimum among all instances. From Table 5.5, it is seen that these instances have the maximum demand on average during the month and even under the perfect information setting amount of overtime and on-call is high.

In order to see the effect of objective function coefficients on the solution of four week schedules, these instances are also solved with objective function coefficients $a = 6$, $b = 2$ and $c = 4$, which reflects a setting where overtime is more valuable than undertime. The results are shown in Table 5.8 and Table 5.9.

Table 5.8 Solution Times, EVPI and VSS Measures for the Instances with 100 Scenarios and Objective Function Coefficients 6, 2, 4

Instance	**Solution Time for Extensive Form**	**Objective Value of Stochastic Solution**	**Over time**	**Under time**	**On-call**	**EEV**	**WS**	**VSS (%)**	**EVPI (%)**
27.1	4.6	820.7	110.0	9.6	61.0	871.6	669.4*	5.9	18.4
27.2	360.7	691.3	81.0	24.8	53.2	715.7	458.3*	3.4	33.7
27.3	5130.8	749.2	93.3	19.5	55.7	768.6	549.5*	2.5	26.6
27.4	647.3	610.3	57.8	43.5	44.6	624.4	310.5*	2.3	49.1
27.5	460.4	692.5	84.6	18.8	54.9	708.3	485.9*	2.2	29.8
27.6	4.2	488.5	21.2	74.6	31.4	492.9	52.0*	0.9	89.0
27.7	4.7	622.6	56.2	49.6	43.5	652.5	276.3*	4.6	55.6
27.8	47.4	1349.2	195.4	8.4	71.4	1431.6	1254.9*	5.8	7.0
27.9	5.0	1757.0	258.9	15.5	70.9	1801.3	1663.0	2.5	5.4
Avg.	740.6	864.6	106.5	29.4	54.1	896.3	635.5	3.3	35.0
Max	5130.8	1757.0	258.9	74.6	71.4	1801.0	1663.0	5.9	89.0

According to Table 5.8, the following inferences are made:

- While for some instances the expected values of perfect information increase in comparison to the first setting, for some instances these values decrease.
- While for some instances the values of the stochastic solutions increase, for some instances the values of the stochastic solutions decrease. When the instances are investigated, it is observed that the instances in which increase is observed have the lowest demand on average during the month.
- The increase in the cost of overtime results in lower overtime and higher undertime.

Table 5.9 Solution Times, EVPI and VSS Measures for the Instances with 3 Scenarios and Objective Function Coefficients 6, 2, 4

Instance	Solution Time for Extensive Form	Objective Value of Stochastic Solution	Over time	Under time	On-call	EEV	WS	VSS (%)	EVPI (%)
28.1	2.1	930.7	123.7	19.8	54.7	986.3	840.0	5.6	9.7
28.2	7219.0	807.3*	93.2	37.2	50.0	830.7	648.3	2.8	19.7
28.3	7213.0	881.3*	108.0	32.8	51.3	920.3	742.7	4.2	15.7
28.4	2.9	755.7	74.2	57.0	41.3	775.7	519.0	2.6	31.3
28.5	3.5	841.0	101.5	31.8	52.3	860.7	705.7	2.3	16.1
28.6	2.5	630.3	34.3	91.8	28.5	646.0	239.0	2.4	62.1
28.7	4.0	742.7	68.2	64.3	38.2	770.3	463.3	3.6	37.6
28.8	5.2	1412.0	204.2	13.3	66.7	1554.0	1344.3	9.1	4.8
28.9	1.8	1802.0	265.0	18.8	68.3	1905.0	1714.0	5.4	4.9
Avg.	1606.0	978.1	119.1	40.8	50.2	1028.0	801.8	4.2	22.4
Max	7219.0	1802.0	265.0	91.8	68.3	1905.0	1714.0	9.1	62.1

According to the Table 5.9, the following inferences are made:

- For all instances the expected values of perfect information decrease in comparison to the first setting. As a result the average of expected values of perfect information decrease.
- While for some instances the values of the stochastic solutions increase, for some instances the value of the stochastic solutions decrease. When the instances are investigated, it is observed that the instances, in which increase is observed, have the lowest demand on average during the month.
- The average of the values of the stochastic solutions decrease in comparison to the first setting.
- The increase in the cost of overtime results in lower overtime and higher undertime. The impact of changing the objective function coefficients is a little more in instances with 3 scenarios, in which time series structure is reflected, in comparison to instances with 100 scenarios.

Last analysis is made by calculating the value of the stochastic solution by using other heuristic approaches than solving the mean value problem. In these heuristic approaches, we solve a deterministic model by using the first quartile (Q_1), median and the third quartile (Q_3) values instead of the average values. The first quartile value is equal to the value that 25% of the scenarios lie below this value and the third quartile value is equal to the value that 75% of the scenarios lie below this value. The value of the stochastic solution, which is calculated according to the median value, is represented as VSS_{50}, the values of stochastic solution, which are calculated according to $Q_1$and Q_3 values, are represented as VSS_{25} and VSS_{75} in the tables below. Together with the VSS values, values of the obtained solutions are also reported as QV_1, QV_2 and QV_3. Table 5.10 and Table 5.11 show the results with the objective function coefficients of 4, 2, 4 and Table 5.12 and Table 5.13 show the results with the objective function coefficients of 6, 2, 4. The last two columns include average values and maximum values of each measure over 9 instances.

Table 5.10 VSS Measures of Instances with 100 Scenarios According to the Heuristic Approaches

Instance	27.1	27.2	27.3	27.4	27.5	27.6	27.7	27.8	27.9	Avg.	Max
Obj. Val. of Stoc. Sol.	600.2	528.6	562.5	494.2	523.3	444.7	508.6	957.7	1239.0	651.0	1239.0
EEV	642.1	544.2	585.3	506.3	535.8	446.2	524.3	1017.0	1276.0	675.3	1276.0
VSS (%)	**6.5**	**2.9**	**3.9**	**2.4**	**2.3**	**0.3**	**3.0**	**5.8**	**3.0**	**3.4**	**6.5**
Avg. Over-time	117.0	83.6	96.9	60.6	86.4	23.2	59.9	207.0	269.7	111.6	269.7
Avg. Under-time	13.9	26.0	22.1	43.9	20.3	72.0	49.4	14.7	16.4	31.0	72.0
Avg. On-call	59.3	52.8	54.7	44.2	54.6	32.8	43.5	65.2	66.0	52.6	66.0
QV_1	606.6	600.1	581.1	628.1	549.7	666.4	676.8	978.0	1268.0	728.3	1268.0
VSS_{25} (%)	**1.1**	**11.9**	**3.2**	**21.3**	**4.8**	**33.3**	**24.9**	**2.1**	**2.3**	**11.6**	**33.3**
Avg. Over-time	112.5	97.1	102.9	78.3	96.1	49.4	77.9	201.0	273.2	120.9	273.2
Avg. Under-time	8.5	24.5	12.7	53.9	12.0	99.0	67.8	9.2	8.1	32.9	99.0
Avg. On-call	61.4	56.9	59.4	49.5	58.6	36.6	46.9	68.6	71.2	56.6	71.2
QV_2	688.8	562.1	624.6	521.8	564.8	465.6	532.7	1056.0	1300.0	701.8	1300.0
VSS_{50} (%)	**12.9**	**6.0**	**9.9**	**5.3**	**7.3**	**4.5**	**4.5**	**9.3**	**4.7**	**7.2**	**12.9**
Avg. Over-time	123.7	82.7	102.2	61.3	90.3	24.6	61.0	213.1	271.6	114.5	271.6
Avg. Under-time	19.4	28.3	26.8	47.2	23.8	75.8	50.4	19.0	21.4	34.7	75.8
Avg. On-call	58.2	53.1	54.1	43.9	54.1	32.2	43.5	63.5	64.0	51.8	64.0
QV_3	809.9	691.1	775.4	629.0	692.8	486.9	626.6	1161.0	1373.0	805.0	1373.0
VSS_{75} (%)	**25.9**	**23.5**	**27.5**	**21.4**	**24.5**	**8.7**	**18.8**	**17.5**	**9.8**	**19.7**	**27.5**
Avg. Over-time	137.8	101.6	122.2	75.4	107.2	20.8	66.6	227.7	280.1	126.6	280.1
Avg. Under-time	36.9	46.5	46.3	61.4	40.5	86.2	70.5	32.7	32.5	50.4	86.2
Avg. On-call	55.5	49.3	50.6	41.0	51.0	29.4	39.0	59.5	61.7	48.6	61.7

According to Table 5.10, the following inferences are made:

- When the average values of the solutions obtained by these methods are considered, it is observed that the lowest value is obtained by solving the expected value problem.
- From Table 5.5, it can be seen that instance 27.6 and instance 27.7 are the instances with the lowest demand on average during the month, and instance 27.8 and instance 27.9 are the instances with the highest demand on average during the month. Therefore, based on the results given in Table 5.10, we can conclude that considering average values and Q_1 values is the best strategy when average demand is low and high, respectively.

According to Table 5.11, the same inferences are made for the setting where we have 3 scenarios.

As can be observed from Table 5.12 and Table 5.13, the relative performance of the heuristic methods remains the same under the setting where overtime cost is higher.

Table 5.11 VSS Measures of Instances with 3 Scenarios According to the Heuristic Approaches

Instance	28.1	28.2	28.3	28.4	28.5	28.6	28.7	28.8	28.9	Avg.	Max
Obj. Value of Stoc. Sol.	683.0	618.8*	664.9	606.7	637.0	558.7	602.0	1003.3	1269.7	738.2	1269.7
EEV	740.3	636.3	696.4*	622.7	654.3	562.7	620.7	1113.3	1357.0	778.2	1357.0
VSS (%)	**7.7**	**2.8**	**4.5**	**2.6**	**2.7**	**0.7**	**3.0**	**9.9**	**6.4**	**4.5**	**9.9**
Avg. Over-time	127.7	95.2	111.7	75.8	103.2	38.3	73.8	219.2	281.0	125.1	281.0
Avg. Under-time	27.7	38.0	35.7	58.3	33.0	87.7	60.5	27.3	25.3	43.7	87.7
Avg. On-call	59.5	51.8	53.5	43.0	54.8	29.3	41.7	63.7	65.8	51.5	65.8
QV_1	745.0	738.3	769.7	828.0	708.3	881.0	801.3	1076.7	1308.7	873.0	1308.7
VSS_{25}(%)	**8.3**	**18.8**	**13.6**	**26.7**	**10.0**	**36.5**	**24.9**	**6.8**	**3.0**	**16.5**	**36.5**
Avg. Over-time	141.0	120.7	134.8	111.5	124.7	75.5	98.2	218.3	279.3	144.9	279.3
Avg. Under-time	18.7	39.5	32.5	73.0	27.2	126.3	80.5	17.8	12.2	47.5	126.3
Avg. On-call	53.2	48.8	50.2	45.0	50.5	36.8	43.3	66.0	71.3	51.7	71.3
QV_2	740.3	965.7	696.4*	622.7	654.3	562.7	620.7	1113.3	1357.0	829.6	1357.0
VSS_{50}(%)	**7.7**	**2.8**	**4.5**	**2.6**	**2.7**	**0.7**	**3.0**	**9.9**	**6.4**	**4.5**	**9.9**
Avg. Over-time	127.7	150.8	111.7	75.8	103.2	38.3	73.8	219.2	281.0	133.7	281.0
Avg. Under-time	27.7	64.3	35.7	58.3	33.0	87.7	60.5	27.3	25.3	48.0	87.7
Avg. On-call	59.5	52.5	53.5	43.0	54.8	29.3	41.7	63.7	65.8	51.3	65.8
QV_3	983.7	851.3	930.7	820.7	871.0	684.7	815.0	1280.3	1499.7	970.8	1499.7
VSS_{75} (%)	**30.6**	**27.3**	**28.6**	**26.1**	**26.9**	**18.4**	**26.1**	**21.6**	**15.3**	**24.6**	**30.6**
Avg. Over-time	157.8	119.2	137.3	95.3	128.5	38.8	84.0	240.7	295.3	144.1	295.3
Avg. Under-time	60.5	68.8	69.2	87.2	63.7	115.5	98.2	49.7	49.0	73.5	115.5
Avg. On-call	55.2	49.7	52.3	45.3	51.2	33.7	43.2	59.5	61.2	50.1	61.2

Table 5.12 VSS Measures of Instances with 100 Scenarios and Objective Function Coefficients of 6, 2, 4 According to the Heuristic Approaches

Instance	27.1	27.2	27.3	27.4	27.5	27.6	27.7	27.8	27.9	Avg.	Max
Obj. Value of Stoc. Sol.	820.7	691.3	749.2	610.3	692.5	488.5	622.6	1349.2	1757.0	864.6	1757.0
EEV	871.6	715.7	768.6	624.4	708.3	492.9	652.5	1431.6	1801.3	896.3	1801.0
VSS (%)	**5.9**	**3.4**	**2.5**	**2.3**	**2.2**	**0.9**	**4.6**	**5.8**	**2.5**	**3.3**	**5.9**
Avg. Over-time	116.5	84.3	95.5	60.2	86.4	23.2	61.0	207.3	267.4	111.3	267.4
Avg. Under-time	13.5	26.3	21.3	43.7	20.2	72.0	50.0	14.6	15.5	30.8	72.0
Avg. On-call	59.4	52.5	55.4	44.4	54.4	32.7	43.1	64.8	67.4	52.7	67.4
QV_1	836.0	781.9	787.4	818.6	742.6	765.1	832.7	1380.1	1815.2	973.3	1815.0
VSS_{25} (%)	**1.8**	**11.6**	**4.9**	**25.5**	**6.7**	**36.2**	**25.2**	**2.2**	**3.2**	**13.0**	**36.2**
Avg. Over-time	113.0	96.1	102.9	81.4	96.3	49.4	77.9	201.0	273.4	121.3	273.4
Avg. Under-time	8.9	23.1	12.8	57.6	12.0	99.0	67.8	9.2	8.2	33.2	99.0
Avg. On-call	61.3	56.4	59.4	50.1	58.4	36.6	46.9	68.6	71.1	56.5	71.1
QV_2	954.4	723.3	831.6	650.6	747.6	512.5	663.2	1461.9	1848.9	932.7	1849.0
VSS_{50} (%)	**14.0**	**4.4**	**9.9**	**6.2**	**7.4**	**4.7**	**6.1**	**7.7**	**5.0**	**7.3**	**14.0**
Avg. Over-time	125.8	84.6	102.8	62.2	90.6	24.7	62.2	210.9	272.3	115.1	272.3
Avg. Under-time	21.1	27.3	26.9	47.7	24.0	75.3	51.0	17.3	21.9	34.7	75.3
Avg. On-call	57.7	53.1	53.6	43.4	54.1	31.6	42.8	63.9	63.8	51.6	63.9
QV_3	1131.0	926.1	982.6	747.9	909.5	527.9	733.4	1622.3	1933.7	1057.0	1934.0
VSS_{75} (%)	**27.5**	**25.4**	**23.8**	**18.4**	**23.9**	**7.5**	**15.1**	**16.8**	**9.1**	**18.6**	**27.5**
Avg. Over-time	143.1	104.8	117.8	71.8	107.0	20.9	64.2	228.1	280.1	126.4	280.1
Avg. Under-time	41.0	49.7	43.0	58.5	41.0	86.0	67.4	33.5	32.5	50.3	86.0
Avg. On-call	54.3	49.2	51.8	41.5	51.7	29.1	39.3	59.9	61.7	48.7	61.7

Table 5.13 VSS Measures of Instances with 3 Scenarios and Objective Function Coefficients of 6, 2, 4 According to the Heuristic Approaches

Instance	27.1	27.2	27.3	27.4	27.5	27.6	27.7	27.8	27.9	Avg.	Max
Obj. Value of Stoc. Sol.	930.7	807.3*	881.3*	755.7	841.0	630.3	742.7	1412.0	1802.0	978.1	1802
EEV	986.3	830.7	920.3	775.7	860.7	646.0	770.3	1554.0	1905.0	1028	1905
VSS (%)	**5.6**	**2.8**	**4.2**	**2.6**	**2.3**	**2.4**	**3.6**	**9.1**	**5.4**	**4.2**	**9.1**
Avg. Over-time	127.3	95.2	111.3	76.2	103.2	39.0	73.8	219.7	280.0	125.1	280
Avg. Under-time	26.3	38.7	36.0	58.3	33.0	88.3	60.8	27.7	23.7	43.7	88.3
Avg. On-call	58.5	52.5	54.2	42.7	54.8	29.3	42.0	64.0	65.2	51.5	65.2
QV_1	1011.0	1007.7	1016.0	1112.7	957.6	1146.0	1015.0	1471.0	1901.0	1182	1901
VSS_{25} (%)	**7.9**	**19.9**	**13.3**	**32.1**	**12.2**	**45.0**	**26.9**	**4.0**	**5.2**	**18.5**	**45.0**
Avg. Over-time	139.7	123.7	132.5	117.2	124.7	86.5	100.8	214.5	283.2	147.0	283.2
Avg. Under-time	16.8	42.2	30.2	79.5	27.2	138.0	81.7	13.3	15.2	49.3	138.0
Avg. On-call	52.7	48.5	50.2	45.8	50.5	37.5	41.8	65.3	70.5	51.4	70.5
QV_2	986.3	1267.3	920.3	775.7	860.7	646.0	770.3	1554.0	1905.0	1076	1905
VSS_{50} (%)	**5.6**	**2.8**	**4.2**	**2.6**	**2.3**	**2.4**	**3.6**	**9.1**	**5.4**	**4.2**	**9.1**
Avg. Over-time	127.3	150..8	11.3	76.2	103.2	39.0	73.8	219.2	280.0	120.1	280.0
Avg. Under-time	26.3	64.3	36.0	58.3	33.0	88.3	60.8	27.7	23.7	46.5	88.3
Avg. On-call	58.5	52.5	54.2	42.7	54.8	29.3	42.0	64.0	65.2	51.5	65.2
QV_3	1274.0	1102.3	1163.7	979.7	1128.0	742.0	958.0	1762.0	2112.0	1247	2112
VSS_{75} (%)	**27.0**	**26.8**	**24.3**	**22.9**	**25.4**	**15.1**	**22.5**	**19.9**	**14.7**	**22.0**	**27.0**
Avg. Over-time	155.0	119.8	133.7	92.7	128.5	37.0	82.5	240.7	297.3	143.0	297.3
Avg. Under-time	58.2	70.5	65.0	83.7	63.7	113.3	95.0	49.7	51.3	72.3	113.3
Avg. On-call	55.7	50.7	52.3	44.5	51.2	33.3	41.5	59.5	61.5	50.0	61.5

CHAPTER 6

CONCLUSION

In this thesis, the scheduling and rescheduling of nurses in an intensive care unit under demand uncertainty is studied. We particularly consider the Cardiovascular Surgery and General intensive care units of a private hospital located in Ankara.

Demand is represented in terms of the required number of nurses. The required numbers of nurses during 9 months are available as the historical data. Time series analysis is used to forecast the future demand. According to our computations, AR(1) model is selected as the most appropriate forecasting model. Scenarios are generated based on the 95% confidence intervals found by the forecasts with AR(1) model.

The problem is modeled as a two-stage stochastic programming model. In the first stage, assignments of nurses to shifts as a normal working nurse and as an on-call nurse are made and the off-days for each nurse are settled. In the second stage, the decision of calling on-call nurses is made and the required amount of overtime or undertime is determined.

Extensive form and L-shaped method are used as the main solution approaches. Since the second stage problem is feasible under every feasible first stage solution, we only use optimality cuts in our L-shaped method. Due to the nice structure of our second stage problem, optimality subproblems are solved by the usage of closed form of dual solutions without solving a LP. In addition to this, a lower bound is added to the master problem. In our computational experiments, it is observed that using the closed form of dual solution and adding the lower bound improves the solution performance in general.

Multi-cut approach for the L-shaped method is also tested and it is observed that the single-cut approach outperforms the multi-cut approach.

In order to improve the solution performance of the L-shaped method, the L-shaped based branch-and-cut method, in which a single search tree is used, is applied. This approach is applied by the usage of lazy constraint callback feature of CPLEX.

Finally, the uncertainty related measures are computed. In the computation of the value of the stochastic solution, different heuristic approaches are used. In these approaches, the values of the stochastic solutions are computed according to the median, the first quartile and the second quartile values. It is observed that considering average values and the first quartile values is the best strategy when average demand is low and high, respectively.

The future studies could be extended to cover the question of which nurse will be assigned as an overtime or undertime nurse besides the decision of amount of overtime and undertime. In addition to this, seasonality on demand could be included in demand forecast with the analysis of the larger historical data.

REFERENCES

[1] S. Arslanhan Memiş, Son Dönemdeki Sağlık Harcamalarının Analizi, Değerlendirme Notu, TEPAV,2012.

[2] P. Punnakitikashem.Integrated Nurse Staffing and Assignment under Uncertainty.PhD thesis, The University of Texas at Arlington, Department of Industrial and Manufacturing Systems Engineering, 2007.

[3] G. B. Dantzig. Linear Programming Under Uncertainty. Management Science, Vol. 1, 197-206, 1955.

[4] J. R. Birge,& F. Louveaux. Introduction to Stochastic Programming.Springer, USA, 1997.

[5] S. Dedu,& F. Serban. Stochastic Optimization Using Interval Analysis with Applications to Portfolio Selection.Journal of Applied Quantitative Methods, 10(1), 30-35, 2015.

[6] H. Madsen. Time Series Analysis. Chapman & Hall/CRC, USA, 2008.

[7] C. Brooks. Introductory Econometrics for Finance. Cambridge University Press, United Kingdom, 2014.

[8] H. Acquah. Comparison of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) in Selection of an Asymmetric Price Relationship. Journal of Development and Agricultural Economics, Vol.2(1), pp 001-006, 2010.

[9] Engineering Statistics Handbook, Introduction to Time Series Analysis, http://www.itl.nist.gov/div898/handbook/pmc/section4/pmc4.htm, last visited on August 2015.

[10] S. T. Rachev, M. Höchstötter, F. J. Fabozzi, & S. M. Focardi.Probability and Statistics for Finance. John Wiley & Sons, Inc., Canada, 2010.

[11] MathWorks, Autocorrelation and Partial Autocorrelation, http://www.mathworks.com/help/econ/autocorrelation-and-partialautocorrelation.html, last visited on August 2015.

[12] M. Sevüktekin, & M. Nargeleçekenler. Zaman Serileri Analizi. Nobel Yayın Dağıtım, Ankara, 2005.

[13] S. Weggemans. Modeling Patients Flow Through a Hospital Using ARIMA Theory and Markov Theory. MS Assignment. University of Twente, Department of Industrial Engineering and Management, 2012.

[14] D. P. Morton, & E. Popova.A Bayesian Stochastic Programming Approach to an Employee Scheduling Problem.IEE Transactions 36, 155-167, 2004.

[15] H.K. Alfares. A Simulation Approach for Stochastic Employee Days-Off Scheduling.International Journal of Modeling and Simulation, Vol. 27, No. 1, 2007.

[16] Campbell, Gerard. A two-stage stochastic program for scheduling and allocating cross-trained workers. Journal of the OperationalResearch Society (2011) 62, 1038–1047. doi:10.1057/jors.2010.16

[17] R. Hall. Handbook of Healthcare System Scheduling. USA. Springer. 2012.

[18] R. Hall. Patient Flow: Reducing Delay in Healthcare Delivery. USA. 2006.

[19] B. Cheang, H. Lim, & B. Rodrigues. Nurse Rostering Problems – a Bibliographic Survey. European Journal of Operational Research, 151, 447-460, 2003.

[20] E.P.C. Kao,& M. Queyranne. Budgeting costs of nursing in a hospital. Management Science, 31(5), 608–621, 1985.

[21] L. Tein, & R. Ramli.Recent Advancements of Nurse Scheduling Models and a Potential Path.6th IMT-GT Conference on Mathematics, Statistics and Applications (ICMSA), Malaysia, 2010.

[22] A. Clark, P. Moule, A. Topping, & M. Serpell. Rescheduling Nursing Shifts: Scoping the Challenge and Examining the Potential of Mathematical Model Based Tools.2012

[23] A.T. Ernst, H. Jiang, M. Krishnamoorthy, & D. Sier. Staff Scheduling and Rostering: A Review of Applications, Methods and Models. European Journal of Operational Research, 153, 3-27, 2004.

[24] J. Belien, & E. Demeulemeester. A Branch and Price Approach for Integrating Nurse and Surgery Scheduling.European Journal of Operational Research, 189, 652-668, 2008.

[25] A. Mobasher, G. Lim , J. F. Bard & V. Jordan. Daily Scheduling of Nurses in Operating Suites, IIE Transactions on Healthcare Systems Engineering, 1:4, 232-246, 2011.

[26] J. F. Bard and H.W. Purnomo. Real-time scheduling for nurses in response to demand fluctuations and personnel shortages. In M.A. Trick and E.K. Burke, editors, Proceedings of the 5th International Conference on the Practice and Theory of Automated Timetabling, pages 67–87, Pittsburgh, PA, 2004.

[27] C. A. Glass, & R. A. Knight. The Nurse Rostering Problem: A Critical Appraisal of the Problem Structure. European Journal of Operational Research, 202, 379-389, 2010.

[28] E. Atmaca, C. Pehlivan, C. B. Aydoğdu, & M. Yakıcı.Hemşire Çizelgeleme Problemi ve Uygulaması. Erciyes Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 28(4), 351-358, 2012.

[29] P. Punnakitikashem, J. M. Rosenberger, and D. B. Behan. Stochastic programming for nurse assignment. To appear in Computational Optimization and Applications, 2006.

[30] K. Kim, Predictive Modeling and Stochastic Optimization in Healthcare. PhD Thesis, Field of Industrial Engineering and Management Sciences, Northwestern University, 2014.

[31] L. M. Schweigler, J. S. Desmond, M. L. McCarthy, K. J. Bukowski, E. L. Ionides, & J. G. Younger. Forecasting Models of Emergency Department Crowd.Academic Emergency Medicine,16, 301–308, 2009.

[32] S. S. Jones, A. Thomas, S. Evans, S.J. Welch, P. J. Haug, & G. L. Snow. Forecasting Daily Patient Volumes in the Emergency Department. Academic Emergency Medicine, 15, 159-170, 2008.

[33] H.J. Kam, J. O. Sung, & R. W. Park. Prediction of Daily Patient Numbers for a Regional emergency Medical Center Using Time Series Analysis. Healthcare Informatics Research, 16(3), 158-165, 2010.

[34] E. P. Kao, & G. G. Tung. Aggregate Nursing Requirement Planning in a Public Health Care Delivery System. Socio-Econ.Pla.Sci., Vol. 15, pp 119-127, 1981.

[35] Registered Nurses Association of Ontario (RNAO). Nursing Workload Measurement Systems- A Discussion of the Issues. 2005.

[36] S. Koç. Geçmişten Günümüze Hemşire Kadro Planlama. Acıbadem Sağlık Grubu.

[37] K. G. Padilha, R. M. C. Sousa., P. C. Garcia, S. T. Bento, E. M. Finardi, & R. H. K. Hatarashi. Intensive and Critical Care Nursing, 26, 108-113, 2010.

[38] R. Lin, M.Y. Sir, E. Şişikoğlu, K. Pasupathy, & L. M. Steege. Optimal Nurse Scheduling Based on Quantitative Models of Work- Related Fatigue. IIE Transactions on Healthcare Systems Engineering, 3, 23-28, 2013.

APPENDIX A

EXAMPLES OF THE MONTHLY SCHEDULES MADE BY HEAD NURSE MANUALLY

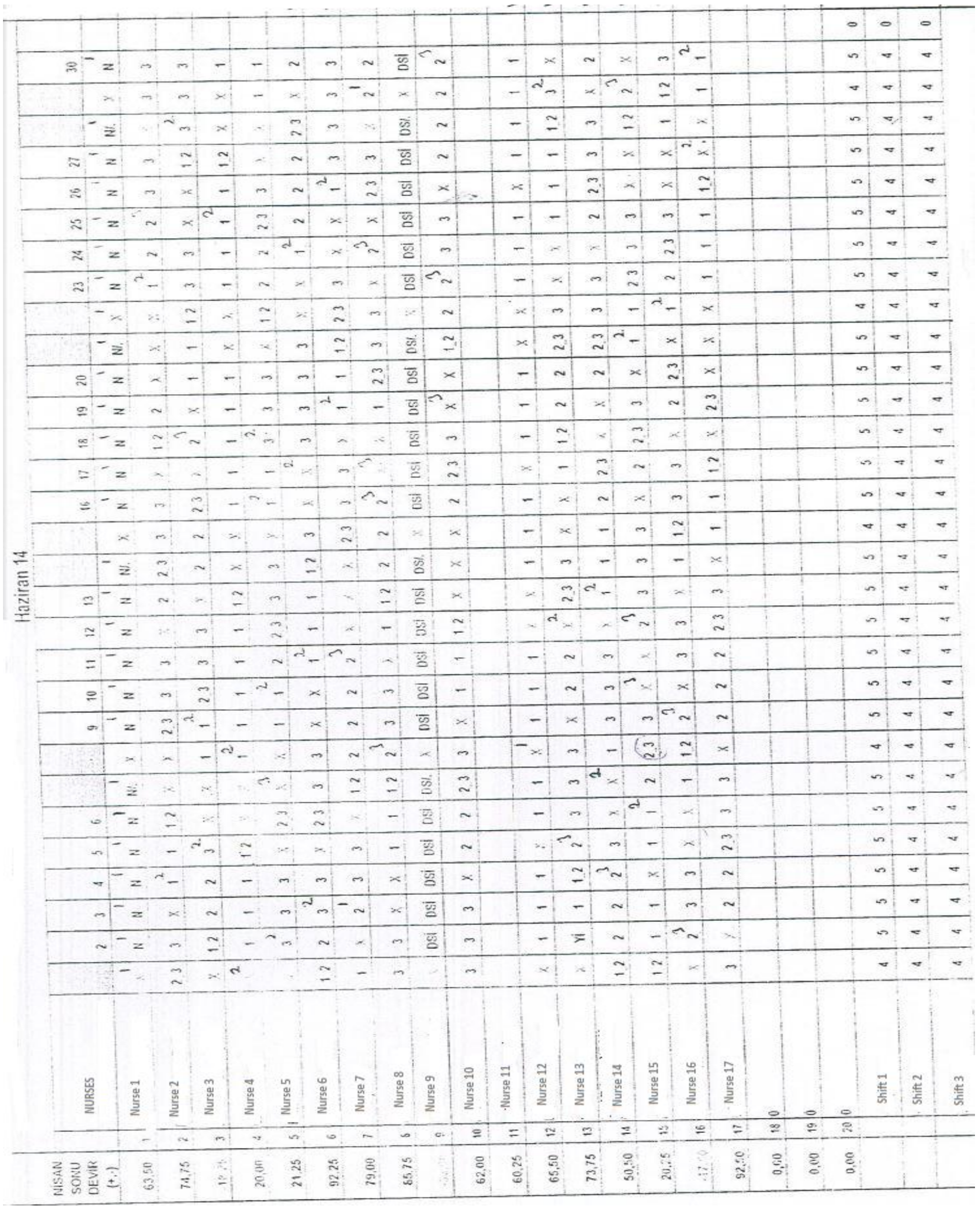


Figure A.1 An Example of the Initial Monthly Schedule

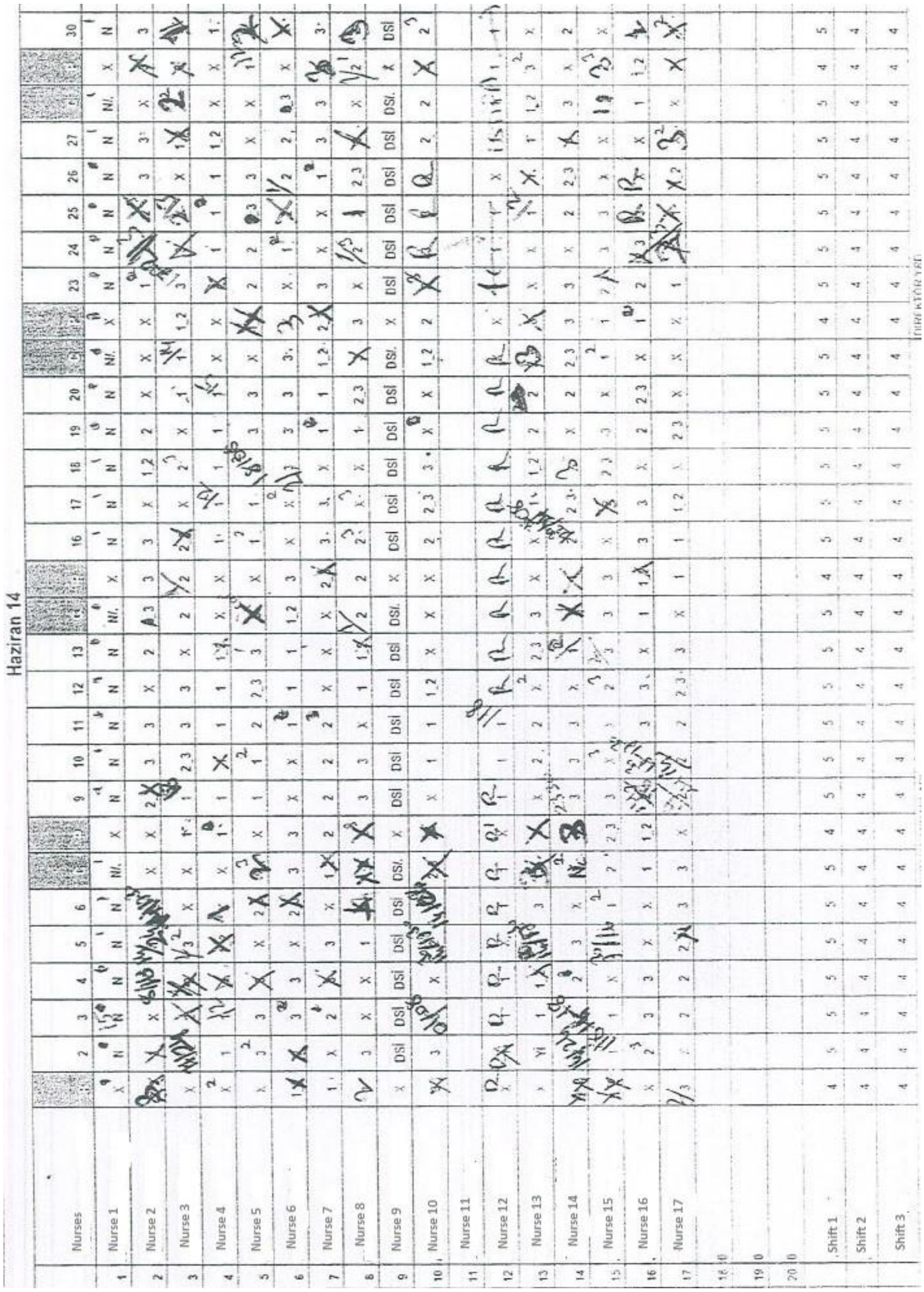


Figure A.2 An Example of Schedule at the End of the Month

APPENDIX B

UNIT ROOT TEST RESULTS

Null Hypothesis: NURSE_COUNT has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.042117	0.0000
Test critical values:	1% level	-3.454534	
	5% level	-2.872081	
	10% level	-2.572460	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(NURSE_COUNT)
Method: Least Squares
Date: 09/15/15 Time: 17:27
Sample (adjusted): 2 270
Included observations: 269 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
NURSE_COUNT(-1)	-0.174926	0.034693	-5.042117	0.0000
C	0.887171	0.194858	4.552918	0.0000

R-squared	0.086939	Mean dependent var	-0.003717
Adjusted R-squared	0.083519	S.D. dependent var	1.407597
S.E. of regression	1.347535	Akaike info criterion	3.441838
Sum squared resid	484.8320	Schwarz criterion	3.468565
Log likelihood	-460.9272	Hannan-Quinn criter.	3.452571
F-statistic	25.42294	Durbin-Watson stat	1.957857
Prob(F-statistic)	0.000001		

Figure B.1 Unit Root Test Results Included Intercept

Null Hypothesis: NURSE_COUNT has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.032603	0.0002
Test critical values:	1% level	-3.992540	
	5% level	-3.426619	
	10% level	-3.136553	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(NURSE_COUNT)
Method: Least Squares
Date: 09/15/15 Time: 17:28
Sample (adjusted): 2 270
Included observations: 269 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
NURSE_COUNT(-1)	-0.175722	0.034917	-5.032603	0.0000
C	0.857200	0.232684	3.683972	0.0003
@TREND("1")	0.000252	0.001065	0.236663	0.8131

R-squared	0.087131	Mean dependent var	-0.003717
Adjusted R-squared	0.080267	S.D. dependent var	1.407597
S.E. of regression	1.349923	Akaike info criterion	3.449062
Sum squared resid	484.7300	Schwarz criterion	3.489152
Log likelihood	-460.8989	Hannan-Quinn criter.	3.465163
F-statistic	12.69453	Durbin-Watson stat	1.956718
Prob(F-statistic)	0.000005		

Figure B.2 Unit Root Test Results Included Intercept and Trend

APPENDIX C

FORECASTED MODELS

Dependent Variable: NURSE_COUNT
Method: Least Squares
Date: 09/14/15 Time: 10:20
Sample (adjusted): 2 270
Included observations: 269 after adjustments
Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.071685	0.469706	10.79756	0.0000
AR(1)	0.825074	0.034693	23.78214	0.0000

R-squared	0.679314	Mean dependent var	5.089219
Adjusted R-squared	0.678113	S.D. dependent var	2.375135
S.E. of regression	1.347535	Akaike info criterion	3.441838
Sum squared resid	484.8320	Schwarz criterion	3.468565
Log likelihood	-460.9272	Hannan-Quinn criter.	3.452571
F-statistic	565.5900	Durbin-Watson stat	1.957857
Prob(F-statistic)	0.000000		

Inverted AR Roots	.83

Figure C.1 Forecasted AR(1) Model for "Nurse_Count" Series

Dependent Variable: NURSE_COUNT
Method: Least Squares
Date: 09/14/15 Time: 10:27
Sample (adjusted): 3 270
Included observations: 268 after adjustments
Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.055722	0.462338	10.93511	0.0000
AR(1)	0.845326	0.061378	13.77237	0.0000
AR(2)	-0.023894	0.061444	-0.388878	0.6977

R-squared	0.679934	Mean dependent var	5.089552
Adjusted R-squared	0.677518	S.D. dependent var	2.379572
S.E. of regression	1.351299	Akaike info criterion	3.451141
Sum squared resid	483.8927	Schwarz criterion	3.491339
Log likelihood	-459.4529	Hannan-Quinn criter.	3.467287
F-statistic	281.4766	Durbin-Watson stat	1.997603
Prob(F-statistic)	0.000000		

Inverted AR Roots	.82	.03

Figure C.2 Forecasted AR(2) Model for "Nurse_Count" Series

Dependent Variable: NURSE_COUNT
Method: Least Squares
Date: 09/14/15 Time: 10:30
Sample: 1 270
Included observations: 270
Convergence achieved after 11 iterations
MA Backcast: 0

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.081166	0.175254	28.99319	0.0000
MA(1)	0.708676	0.043334	16.35389	0.0000

R-squared	0.496297	Mean dependent var	5.085185
Adjusted R-squared	0.494418	S.D. dependent var	2.371642
S.E. of regression	1.686340	Akaike info criterion	3.890378
Sum squared resid	762.1231	Schwarz criterion	3.917033
Log likelihood	-523.2010	Hannan-Quinn criter.	3.901081
F-statistic	264.0596	Durbin-Watson stat	1.312475
Prob(F-statistic)	0.000000		

Inverted MA Roots	-.71

Figure C.3 Forecasted MA(1) Model for "Nurse_Count" Series

Dependent Variable: NURSE_COUNT
Method: Least Squares
Date: 09/14/15 Time: 10:32
Sample: 1 270
Included observations: 270
Convergence achieved after 19 iterations
MA Backcast: -1 0

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.081583	0.207684	24.46780	0.0000
MA(1)	0.822148	0.054901	14.97506	0.0000
MA(2)	0.440838	0.054902	8.029581	0.0000

R-squared	0.597994	Mean dependent var	5.085185
Adjusted R-squared	0.594983	S.D. dependent var	2.371642
S.E. of regression	1.509336	Akaike info criterion	3.672266
Sum squared resid	608.2517	Schwarz criterion	3.712249
Log likelihood	-492.7559	Hannan-Quinn criter.	3.688321
F-statistic	198.5845	Durbin-Watson stat	1.682726
Prob(F-statistic)	0.000000		

Inverted MA Roots	-.41+.52i	-.41-.52i

Figure C.4 Forecasted MA(2) Model for “Nurse_Count” Series

Dependent Variable: NURSE_COUNT
Method: Least Squares
Date: 09/14/15 Time: 10:33
Sample (adjusted): 2 270
Included observations: 269 after adjustments
Convergence achieved after 7 iterations
MA Backcast: 1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.073255	0.457092	11.09899	0.0000
AR(1)	0.813562	0.043361	18.76268	0.0000
MA(1)	0.035745	0.074309	0.481036	0.6309

R-squared	0.679547	Mean dependent var	5.089219
Adjusted R-squared	0.677138	S.D. dependent var	2.375135
S.E. of regression	1.349575	Akaike info criterion	3.448546
Sum squared resid	484.4795	Schwarz criterion	3.488635
Log likelihood	-460.8294	Hannan-Quinn criter.	3.464646
F-statistic	282.0376	Durbin-Watson stat	2.002319
Prob(F-statistic)	0.000000		

Inverted AR Roots	.81
Inverted MA Roots	-.04

Figure C.5 Forecasted ARMA(1,1) Model for “Nurse_Count” Series

Dependent Variable: NURSE_COUNT
Method: Least Squares
Date: 09/14/15 Time: 10:35
Sample (adjusted): 3 270
Included observations: 268 after adjustments
Convergence achieved after 10 iterations
MA Backcast: 2

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.055386	0.464805	10.87636	0.0000
AR(1)	-0.043343	0.134794	-0.321548	0.7480
AR(2)	0.705077	0.119994	5.875952	0.0000
MA(1)	0.908738	0.111373	8.159393	0.0000

R-squared	0.682435	Mean dependent var	5.089552
Adjusted R-squared	0.678827	S.D. dependent var	2.379572
S.E. of regression	1.348555	Akaike info criterion	3.450757
Sum squared resid	480.1103	Schwarz criterion	3.504354
Log likelihood	-458.4014	Hannan-Quinn criter.	3.472284
F-statistic	189.1090	Durbin-Watson stat	2.024546
Prob(F-statistic)	0.000000		

Inverted AR Roots	.82	-.86
Inverted MA Roots	-.91	

Figure C.6 Forecasted ARMA (2,1) Model for "Nurse_Count" Series

Dependent Variable: NURSE_COUNT
Method: Least Squares
Date: 09/14/15 Time: 10:36
Sample (adjusted): 2 270
Included observations: 269 after adjustments
Convergence achieved after 10 iterations
MA Backcast: 0 1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.057453	0.532165	9.503541	0.0000
AR(1)	0.871690	0.044210	19.71695	0.0000
MA(1)	-0.041987	0.077589	-0.541151	0.5889
MA(2)	-0.128356	0.073575	-1.744572	0.0822

R-squared	0.681702	Mean dependent var	5.089219
Adjusted R-squared	0.678099	S.D. dependent var	2.375135
S.E. of regression	1.347565	Akaike info criterion	3.449233
Sum squared resid	481.2215	Schwarz criterion	3.502686
Log likelihood	-459.9219	Hannan-Quinn criter.	3.470700
F-statistic	189.1844	Durbin-Watson stat	1.979746
Prob(F-statistic)	0.000000		

Inverted AR Roots	.87	
Inverted MA Roots	.38	-.34

Figure C.7 Forecasted ARMA(1,2) Model for “Nurse_Count” Series

Dependent Variable: NURSE_COUNT
Method: Least Squares
Date: 09/14/15 Time: 10:38
Sample (adjusted): 3 270
Included observations: 268 after adjustments
Convergence achieved after 10 iterations
MA Backcast: 1 2

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.054143	0.476326	10.61068	0.0000
AR(1)	0.007282	0.213344	0.034135	0.9728
AR(2)	0.678531	0.162567	4.173866	0.0000
MA(1)	0.844894	0.221773	3.809728	0.0002
MA(2)	-0.031122	0.087837	-0.354318	0.7234

R-squared	0.682561	Mean dependent var	5.089552
Adjusted R-squared	0.677733	S.D. dependent var	2.379572
S.E. of regression	1.350848	Akaike info criterion	3.457822
Sum squared resid	479.9197	Schwarz criterion	3.524818
Log likelihood	-458.3482	Hannan-Quinn criter.	3.484731
F-statistic	141.3767	Durbin-Watson stat	2.000178
Prob(F-statistic)	0.000000		

Inverted AR Roots	.83	-.82
Inverted MA Roots	.04	-.88

Figure C.8 Forecasted ARMA(2,2) Model for “Nurse_Count” Series