Hemoglobin Modeling and Simulation for Anemia Management in Chronic Kidney Disease

by

Nirwair Singh Bajwa

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Process Control

Department of Chemical and Materials Engineering

University of Alberta

©Nirwair Singh Bajwa, 2018

Abstract

Chronic Kidney Disease (CKD) affects huge number of people in the world. One of the significant side effects of this infection is the inability to maintain the body's red blood cell production (RBC), and along these lines, the mass of a protein called hemoglobin inside the body. The wellbeing of these patients crumbles and over time they become anemic. Recently, exogenous erythropoietin stimulating agents have turned into the standard for treating anemia during CKD. The pharmaceutical works greatly well for what it is intended to do. The issue with this situation is the failure of the doctor's to have the capacity to pick an appropriate dose for every patient. The dosing protocols are not standardized crosswise over hospitals, and huge numbers of the dosings regimens are ineffectively designed. In that capacity, many patients' hemoglobin levels are inadequately controlled. The poor hemoglobin control in CKD patients is well documented in the literature. Automated anemia management based on feedback is one approach to address this issue. A relaible model is very important in the design of the control system.

In this thesis, the objectives are (1) to present an artificial patient simulator developed exclusively based on measurement noise and time-varying parameters in Pharmacokinetics and Pharmacodynamics (PKPD) model, (2) performance assessment of non-linear constrained ARX model (C-ARX), and (3) hemoglobin modeling technique with modified constrained ARX modeling (C-ARX) method utilizing additional measurment such iron saturation and white blood cell (WBC) count. The hemoglobin response modeling methods are compared on a clinical data containing 167 patients. It will be demonstrated that the new modeling method offers better modeling results to the previously developed C-ARX model.

Acknowledgements

Foremost, I would like to thank my research supervisors Dr. Jinfeng Liu and Dr. Zukui Li from the Chemical and Materials Engineering Department at the University of Alberta. Without their supervision, this thesis would not have been possible. Both Dr. Li and Liu have contributed one of their unique aspects to help in the formation of this thesis.

Without Dr. Li, the hemoglobin modeling with additional measurements and model performance assessment work would not have been possible. I esteem his endless wisdom in the field of optimization and express gratitude toward him for offering his insight to me. I am grateful to Dr. Liu's incredible help with respect to each step during my masters. His in-depth insights on control theory was very useful, and he seemed to always come up with unique ideas whenever I had failures with my work. I learned a lot about research in process control during these two years from Dr. Liu and Dr. Li and commitment to me, and I will dependably feel thankful to them.

I am also thankful to the organizations that supported this project with research funding. This incorporates the Graduate Student's Association for the Academic travel award to enable me to present my research at the Chemical Society for Chemical Engineers Conference in Edmonton. I am grateful to the Natural Sciences and Engineering Research Council of Canada for financing this project and express gratitude toward Cybernius Medical Ltd. for their contribution. I am likewise appreciative to Faculty of Graduate Studies for Captain Thomas Farrell Greenhalgh Memorial Graduate Scholarship, I received from them.

I am thankful to Dr. Stevan Dubljevic for the opportunity to become in his class to become teaching assistant in his class. Being a TA allowed me to refresh my basic mathematical concepts and also allowed me to teach future leaders.

Last yet not slightest, I should thank the other graduate students who have con-

tributed essentially to my advance. These incorporate Jayson McAllister, Hossein Shahandeh, Su Liu, Jannatun Nahar, An Zhang, Azzam Hazim, Soumya Sahoo, and Benjamin Decardi-Nelson. I wish them good luck for their brilliant future.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

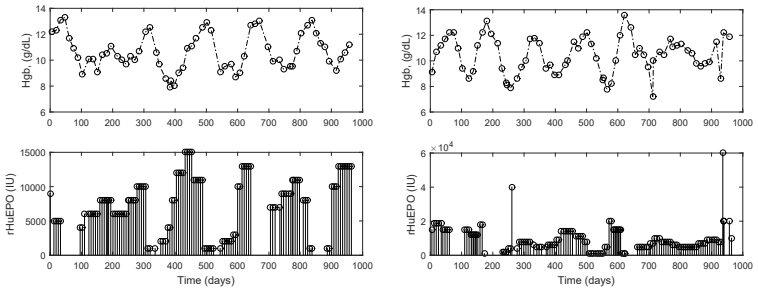
1.1 Motivation

Chronic Kidney Disease (CKD) is assessed to affect about 10% of the world's populace (A. Levey and Coresh, 2007). There are a few phases of CKD that patients can be classified into, with the most extreme stage being categorized as End Stage Renal Disease (ESRD). A huge number of individuals/patients worldwide are classified as ESRD patients, what's more, numerous experience dialysis treatment or kidney transplants subsequently (W. Couser *et al.*, 2011). More than 80% of chronic kidney disease patients that get treatment are in affluent nations that approach all inclusive human services and have huge elderly populaces (V. Jha and Iseki, 2013). The elderly populace is expanding at a high rate in developing countries, for example, China and India and the quantity of interminable kidney patients is expected to increment extremely in the coming years (V. Jha and Iseki, 2013).

One of the significant side effects of CKD is the powerlessness to generate endogenous erythropoietin, which is body's hormone to manage the generation of body's red blood cells (RBCs). RBCs contain hemoglobin, a protein which is essential to human survival. Hemoglobin is incharge for binding to oxygen and conveying it throughout the body tissues and organs. Since without oxygen tissue will die, therefore, maintaining hemoglobin is necessary. When the endogenous generation of erythropoietin drops considerably, these patients suffers with anemia disease, which is described as decreased hemoglobin and RBC mass in the body. During the 1980s researchers understood effect of recombinant human erythropoietin (rHuEPO) that helps manage the creation of RBCs and subsequently hemoglobin. In this period, kidney patients

bearing anemia began experiencing erythropoietin stimulating agent (ESA) therapy. After the initial discovery of rHuEPO, many alternate drugs have been developed to increase RBCs production, examples include epoetin-alfa (EPO), methoxy polyethylene glycol-epoetin beta, darbepoetin-alfa, etc (McAllister, 2017). In this thesis, all analysis is done based on EPO as a stimulating agent (ESA) therapy.

Clinicians have demonstrated that while a low level of hemoglobin moves the patient towards anemia, but a higher level of hemoglobin more than 13.0 g/dL can build the danger of mortality risk for the patients already suffering from CKD (M. Rosner, 2008; Z. Jing *et al.*, 2012; A. Singh *et al.*, 2006). Subsequently, effective techniques are required to decide the proper dosage of erythropoietin stimulating agent to keep hemoglobin within target amount. Numerous Anemia Management Protocols (AMPs) utilized by clinicians, depend on the set of rules made by clinicians which are based on past experiences or retrospective studies. These methods are not consistent between hospitals and often results in incorrect drug dose to the patient. In a previous study (McAllister, 2017), it was found that out of 167 cases about 56% of patients suffered oscillations in hemoglobin because of imprecise drug dose. Figure 1.1 shows two such patients where target hemoglobin control zone is 9.5 to 11 g/dL (McAllister, 2017). These one-size-fits-all convention methods for rHuEPO administration are generally imprecise, tedious, and non-robust and as a result, the patient's hemoglobin levels are often poorly controlled. Therefore, new methods are required to avoid the mortality effect associated with high levels of hemoglobin, at the same time decreasing anemia effects. Assessed by National Kidney Foundation, 10% of the worldwide population is affected by CKD and millions of people die every year due to unaffordable access to treatment (National Kidney Foundation, 2017). Treatment associated with CKD costs the United States 48 billion dollars per year (National Kidney Foundation, 2017) for the treatment of ($< 1\%$) fraction population (P. Damien *et al.*, 2016). Hence, the potential for the opportunity for cost savings becomes important while enhancing the well-being of patients health.



(a) Patient 1 **(b)** Patient 2

Figure 1.1: Actual clinical data showing patients undergoing hemoglobin cycling (McAllister, 2017)

The ideal hemoglobin target is still under discussion (Singh, 2007) yet for the most part lies between 9.5 and 12.0 g/dL for CKD patients. In McAllister, 2017 a range of 9.5 to 11.0 g/dL was used for control zone.

One of the physician's protocol has been shown in Figure 1.3 from one of the hospitals with sensitive information removed. It can be seen that the one-size fit rule protocol is only sensitive to patient's weight and all other guidelines have been ignored. Classical and zone model predictive controller (ZMPC) were compared in McAllister, 2017 and recommended a weighted recursive least square ZMPC that uses a funnel-shaped control zone, Figure 1.2.

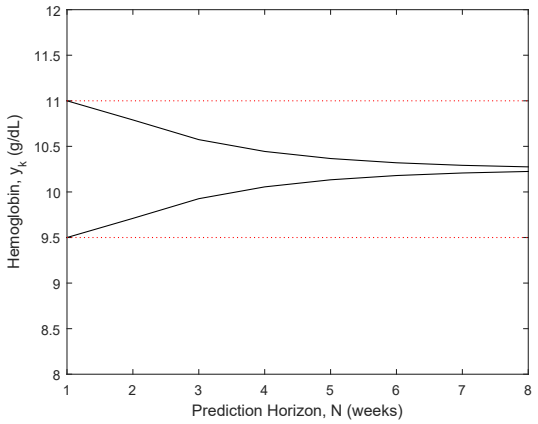


Figure 1.2: Control zone funnel shape Constraint boundaries (McAllister, 2017)

PHYSICIAN ORDERS
EPOETIN ALFA (EPOGEN) IN HEMODIALYSIS
Page 1 of 1

All ☒ orders are active unless lined through and initialed. ❑ Boxes are optional. Check or circle orders as needed.

ALLERGIES: ❑ NKDA ❑ NKFA ❑ Latex ❑ Other: _____

- ☒ **Epoetin alfa[BBW] dose is not to exceed 10,000 units IV 3x/week**
- ☒ **Round Epoetin alfa[BBW] dose to the nearest 1,000 units**
- ☒ **Epoetin alfa[BBW] dose is to be administered during hemodialysis**
- ☒ **Target hemoglobin (Hgb): 9.5-11.0 g/dL**
- ☒ **Dose increase is not to be made more frequently than q 4 weeks unless ordered by a physician**
- ☒ **Dose reduction is not to be made more frequently than q 2 weeks**
- ☒ **If predialysis systolic BP greater than 190 mmHg or diastolic BP greater than 100 mmHg, check BP q 30 minutes**
  - **If SBP less than or equal to 190 mmHg and DBP is less than or equal to 100 mmHg, give Epoetin alfa[BBW] at current dose**
  - **If SBP greater than 190 mmHg or DBP greater than 100 mmHg, hold Epoetin alfa[BBW] dose**
- ☒ ~~**Check Initial CritLine every treatment**~~
  ~~**Hct less than or equal to 33%: Give prescribed Epoetin alfa[BBW] dose**~~
  ~~**Hct greater than 33%: Hold Epoetin alfa[BBW], monitor dialyzer for clotting**~~
- ☒ **Notify attending physician for:**
  - **SBP greater than 190 mmHg and DBP greater than 100 mmHg resulting in Epoetin alfa[BBW] held x 2 months**
  - **Hemoglobin (Hgb) less than 9 g/dL or no response within 1 month**
    - **Assess for iron deficiency, bleeding, infection, and/or inflammatory process**
  - **Decrease in Hemoglobin (Hgb) greater than 1g in 1 month**

Epoetin alfa Initiation Criteria

Hemoglobin (g/dL)	**Initiation** of Epoetin alfa	Lab Monitoring
Less than 10	100 units/kg IV 3x/week	TSAT, ferritin prior to initial dose, q ~~month~~2 weeks
10-10.9	50 units/kg IV 3x/week	TSAT, ferritin prior to initial dose, q ~~month~~2 weeks
Greater than or equal to 11	Do not initiate Epoetin	Hgb and Hct, q ~~month~~2 weeks

Epoetin alfa Titration

Current Hemoglobin (g/dL)	Epoetin alfa **Dose Adjustment**	Lab Monitoring
Less than 9.5	Increase dose by 25% or increase to 100 units/kg IV 3x/week, whichever is higher	Hgb and Hct q 2 weeks
9.5-11.0	Maintain dose	Hgb and Hct q ~~month~~2 weeks
11.1-11.5	Decrease dose by 10%	Hgb and Hct q 2 weeks
Greater than 11.5	Hold dose for 2 weeks then resume when Hgb less than 11.5 at a dose ~~Decrease~~ decreased ~~dose~~ by 20%. If in 2 weeks Hgb is still greater than 11.5 ~~decrease dose by another 10%. If Hgb~~	Hgb and Hct q 2 weeks ~~If Hgb greater than 11.5 for 3 months, then check Hgb and Hct q month~~

Date Time Signature MD Provider Stamp (or MD Print name/ Hospital number)

~~Date: Time Noted by: RN~~

Figure 1.3: Physician's Protocol for dosing Epoetin-alfa

1.2 Contributions

The target of this thesis was to investigate and improve several innovations for the usage of an automatic dosage optimizer for the administration of exogenous drug dosage for patients suffering with anemia in end-stage renal disease. The application could be executed in programming projects to automatically collect past measurements from the patient database, obtain a specific patient model, and recommend an enhanced EPO dosing regimens to the numerous patients undergoing end-stage renal disease treatment. This thesis aims to improve existing modeling techniques and relates mathematical models with biological systems in the human body to make mathematical models more acceptable to medical professionals. This thesis is in continuation of future work identified in McAllister, 2017 and cover three sections. First, model performance assessment for constrained ARX model (C-ARX) developed by (J. Ren *et al.*, 2017). Second, incorporating additional measurements such as Iron Saturation and White Blood Cell Counts (WBCs) in constrained ARX model identification. Third, a designed artificial patient simulator which based on biological systems. Overall, this thesis aims to improve health for CKD patients by helping clinicians to make a better decision with epoetin-alfa (EPO) dosing regimens.

1.3 Thesis outline

Chapter 2 focuses on model performance assessment for constrained ARX model (C-ARX) developed by (J. Ren *et al.*, 2017). This chapter begins with preliminaries for existing mathematical modeling for anemia using EPO, introduction to the problem description of model performance assessment. It follows data pre-processing and then four classification techniques are introduced in detail including Principal Component Analysis (PCA), Partial Least Square - Discriminant Analysis (PLS-DA), Linear Discriminant Analysis (LDA), and combination of PCA-LDA. These classification techniques are used on patients actual clinical data along with modeling parameters identified from C-ARX to analyze the decision whether patients original data is good enough to identify constrained ARX model. This chapter shows that classification the technique developed by combining PCA and LDA is the most effective way to

decide modelling for original data. It will be also shown how classification accuracy for PCA-LDA can be increased by outliers detection from original data. Outlier detection section can be used as guidelines for the patients where C-ARX model will not result into good hemoglobin prediction model and C-ARX model should not be used for those patients until patient's original data return to normal range.

Chapter 3 focuses on future work identified in (McAllister, 2017). This includes incorporating additional measurements such as iron saturation, White Blood Cell Counts (WBCs), and ferritin level. The chapter begins with an introduction to physical significance of additional parameters followed by modeling techniques to update the constrained ARX model developed by developed by J. Ren *et al.*, 2017. Improved model identification was done in three parts. First, by introducing Saturation Iron in the model. Second, by sequentially optimizing with Iron Saturation and WBC, and third by simultaneous optimization with Iron Saturation and WBC. It will be shown that simultaneous optimization or sequential optimization with Iron Saturation, or original constrained ARX model can be used under different conditions to improve model performance.

In chapter 4, an artificial patient simulator is discussed in detail that has been designed based on biological systems by identifying future work in McAllister, 2017. The artificial patient mimic real life disturbances that can happen to CKD patients and have been paired with parameters identified in pharmacokinetics and pharmacodynamics (PKPD) model (Y. Chait *et al.*, 2014). The chapter begins with motivation and interpreting PKPD model parameters, followed by different scenarios that CKD patients suffer. Later, it was discussed that how by manipulating the physical parameters, and by tuning process disturbance and measurement can relate to change in patient health. Actual clinical figures are compared with simulated figures to show how the developed simulator can be explained to professionals in the medical industry.

Chapter 5 identifies some future work and considerations for this project. Most specially is controller performance assessment with the developed artifical patient simulator. Other future research areas identified are hemoglobin response modeling with iron saturation and introducing iron dosage, and actual MCH in Modelling.

Chapter 2

Model Performance Assessment of Constrained ARX model

2.1 Introduction

The aim of this chapter is model performance assessment of constrained ARX model (C-ARX) developed in J. Ren *et al.*, 2017. C-ARX model did not always result into a good model, therefore, model performance assessment of C-ARX was required. There were two data sets used in this assessment. The first data set contains 1-3 years of clinical data for 145 patients and second data set contains 209 patients. Hemoglobin values were typically taken approximately 2 weeks apart, while ESA dosing was done typically once per week. Using these data sets assessment parameters were identified using raw data and C-ARX model. These parameters were then used to explore various classification techniques. The chapter starts with preliminaries required for hemoglobin modeling, followed by problem description and classification techniques.

2.2 Constrained autoregressive with exogenous inputs model (C-ARX)

In J. Ren *et al.*, 2017, an Autoregressive with Exogenous Inputs (ARX) modeling approach was developed. It was shown that the developed C-ARX modeling strategy had better performance than PKPD whilst being less complex at the same time.

The developed C-ARX model has one 'a' parameter and 20 'b_k' parameters. The 'b_k' parameters are constrained to follow a particular shape. The model structure is

presented in Equation (2.1)

$$\begin{aligned} Hgb_{t+1} - Hgb_{ss} =& a_1 \left(Hgb_t - Hgb_{ss}\right) \ldots \\ &+ \sum_{k=1}^{20} b_k \left(EPO_{t-k+1} - EPO_{ss}\right) + e_t \end{aligned} \tag{2.1}$$

where Hgb_{ss} and EPO_{ss} are the steady state hemoglobin and the erythropoietin respectively. Parameter estimation of the above model is done by formulating it as a (MINLP) and is outlined in Equation (2.2).

$$\min \sum_{t=1}^{tf} [Hgb_t - Hgb_{t,actual}]^2 \tag{2.2a}$$

$$s.t. - Kz_k + 0.001 \leq k - t_{peak} \leq K(1 - z_k) \qquad \forall k = 1, \ldots, K \tag{2.2b}$$

$$-M(1 - z_k) \leq \alpha(k-1) - b_k(t_{peak} - 1) \leq M(1 - z_k) \qquad \forall k = 1, \ldots, K \tag{2.2c}$$

$$-Mz_k \leq \alpha \exp^{-\beta(k - t_{peak})} - b_k \leq Mz_k \qquad \forall k = 1, \ldots, K \tag{2.2d}$$

$$7.0 \leq Hgb_t \leq 15.0 \quad \forall t \tag{2.2e}$$

$$7.0 \leq Hgb_{ss} \leq 11.0 \tag{2.2f}$$

$$0.7 \leq a_1 \leq 0.99 \tag{2.2g}$$

$$0 \leq EPO_{ss} \tag{2.2h}$$

$$b_1 = 0 \qquad b_k \geq 0 \quad \forall k = 2, \ldots, K \tag{2.2i}$$

$$b_k \geq 0.1 \quad k = k_{peak} \qquad 1.1 \leq t_{peak} \leq 3.9 \tag{2.2j}$$

$$\alpha \geq 0.1 \qquad \beta \geq 0.05 \qquad z_k \in \{0, 1\} \tag{2.2k}$$

The objective function (Equation (2.2a)) is sum of square errors between actual hemoglobin and proposed hemoglobin level by the model. The Equation (2.2b) enforces the constrain on the binary variable z_k that are introduced to represent whether a time occasion k is equivalent to after the peak time ($z_k = 0$) or before the peak time ($z_k = 1$). Equation (2.2c) constrains model parameters b_k before the peak time and Equation (2.2d) constrains after the peak time. Equation (2.2e) constrains hemoglobin values within reasonable values. Equations (2.2f) and (2.2h) constrain steady state hemoglobin and erythropoietin respectively. Equations (2.2g), (2.2i), (2.2j), and (2.2k) constrain range for model parameters.

2.3 Problem description for model performance assessment

Analysis of Constrained ARX model (C-ARX) developed in (J. Ren *et al.*, 2017) on 145 patients showed that 35 patients did not result into the good model. For these 35 patients, either model dynamics did not capture actual data dynamics, or residuals were too high, or patient did not had enough process excitation to be classified as a good model. Typically for the model performance, one looks at the 1-step prediction residual only (Clements and Hendry, 2005), but residuals always may not accuratly guide regarding good model performance. For instance, if a paitent had no EPO dosages it's 1-step resduals may still be lower, leading to missclassification as a 'good model'. Considering other parameters such as predicting 8-step ahead, checking process excitation, and assessing overall other parameters that could affect model performance is important. This led to the motivation for the work that under what scenarios model developed by C-ARX will not result in good model. There are two objectives as follow:

1. To determine the most significant variables in describing the dynamics of hemoglobin, which is the variable used in EPO dosage determination.

2. To develop a method to assess model performance automatically using the variables determined in the first objective.

2.4 Data Selection and Preprocessing

There was two datasets used. First data set had 167 patients and second data set had 310 patients. After removing patients which did not have enough data points (less than 60 weeks of data) or missing data, the first data set was reduced to 145 patients and second data set to 209 patients. Dataset two (209 patients) was used as training dataset which includes training and cross-validation, whereas data set one (145 patients) was used as a test set for prediction.

For the reason of simplicity, it was assumed that weekly hemoglobin measurements and weekly EPO dosages are available. The clinically ideal Hgb measurement

frequency for CKD patients is 4 times each month (A. Gaweda *et al.*, 2010). Since the available data does not have a weekly measurement, data must be re-sampled.

Generally, Hgb measurement is bi-weekly, but measurement and EPO dosing could vary drastically over the patient's past history. Ordinarily, the dosing was done each week, yet it was normal to see a patient not getting a dosage for a long time, or get numerous measurements around the same time. The patient information utilized had each aggregate medication dosage in worldwide units administered in multiples of thousands. The week by week total of dosages went from 0 IU to as high as 45 000 IU.

As proposed by McAllister, 2017 it is required to use re-sampled data into weekly intervals. Figure 2.1 contains a pictorial case where the first information is appeared in blue and the re-examined information focuses are appear in red (McAllister, 2017). The weekly hemoglobin measurement was estimated by linear interpolation and EPO dosages were clubbed together for weekly dosages by taking the total of the present days EPO dosage and the past 6 days of dosages.

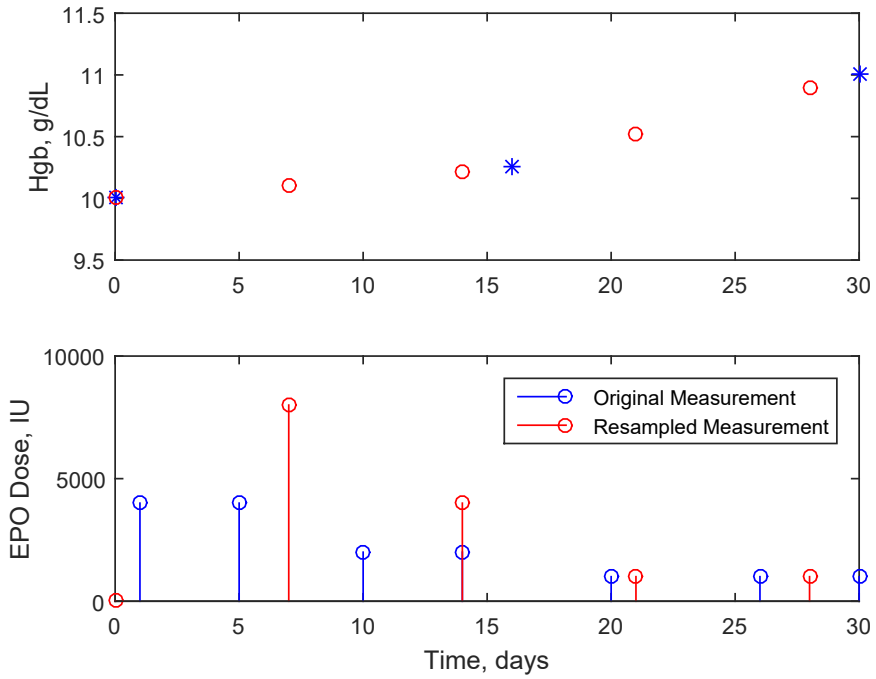


Figure 2.1: Patient Data Resampling Example

2.4.1 Performance related variable selection

Different statistic indices/variables were calculated based on the patients' raw data and the prediction residulas of C-ARX models. This includes 1-step mean, 1-step standard error, 8-step mean, 8-step standard error, 1-step standard deviation, 8-step standard deviation, raw data input signal to noise ratio (SNR), raw data output SNR, average EPO dose, patient weight, sum of b parameters, input-output means and standard deviations and cross-correlations at different lags. After various trials and analysis, following set of variables were found to affect model quality significantly. The following section describes each parameter used:

1. **Mean hemoglobin 1-step prediction error**:

   Hemoglobin 1-step prediction errors are calculated by predicting hemoglobin 1-week ahead using *compare* function in MATLAB. Mean hemoglobin 1-step prediction error is calculated by taking mean of all 1-step prediction errors.

2. **Standard error of hemoglobin 1-step prediction errors**:

   It is the standard deviation of the sampling distribution (hemoglobin 1-step prediction errors) of the mean. The formula for the hemoglobin 1-step prediction errors is:

   $$\sigma_M = \sigma/\sqrt{N} \tag{2.3}$$

   where σ is the standard deviation of all the hemoglobin 1-step prediction errors and N is the sample size (length of weekly data).

3. **Mean hemoglobin 8-step prediction error**: Hemoglobin 8-step prediction errors are calculated by predicting hemoglobin 8-week ahead using *compare* function in MATLAB. Mean hemoglobin 8-step prediction error is calculated by taking mean of all 8-step prediction errors.

4. **Standard error of hemoglobin 8-step prediction errors**: It is the standard deviation of the sampling distribution (hemoglobin 8-step prediction errors) of the mean.The formula for the standard error of the mean is:

   $$\sigma_M = \sigma/\sqrt{N} \tag{2.4}$$

where σ is the standard deviation of the hemoglobin 8-step prediction errors and N is the sample size (length of weekly data).

5. **1-step standard deviation**: It is the standard deviation of all the hemoglobin 1-step prediction errors.

6. **8-step standard deviation**: It is the standard deviation of all the hemoglobin 8-step prediction errors.

7. **Sum of b parameters**: It is the sum of all b parameters. The formula for the sum of b parameters is:

$$sum_b = \sum_{k=1}^{20} b_k \tag{2.5}$$

8. **Raw data input mean**: It is the mean of all raw data EPO dosages for a particular patient.

$$RawInput_m = \frac{1}{N} \sum_{t=1}^{N} EPO_t \tag{2.6}$$

9. **Raw data input standard deviation**: It is the standard deviation of all raw data EPO dosages for a particular patient.

10. **Raw data output mean**: It is the the mean of all raw data Hgb measurements available for a particular patient.

11. **Raw data output standard deviation**: It is the standard deviation of all raw data Hgb measurements available for a particular patient.

12. **Cross correlation function with lag from 1 to 5**: The normalized cross-covariance between input sequence (EPO) and output (Hgb) was estimated with lags from 1 to 5 for Cross Correlation Function (CCF). Since EPO dosage do not have immidiate effect on Hgb and the delay time varies for different patients (Y. Chait *et al.*, 2014), CCF at lag 0 was not considered. The normalized cross-covariance (CCF) between EPO and Hgb was calculated with the following

equation:

$$\gamma_{xy}(l) = E[(EPO(n) - EPO_{ss})(Hgb(n-l) - Hgb_{ss})]$$
$$\rho_{xy}(l) = \frac{\gamma_{xy}(l)}{\sqrt{\sigma^2_{EPO}\sigma^2_{Hgb}}} \tag{2.7}$$

where E is the expectation, $\rho_{xy}(l)$ is CCF at different lags, $\gamma_{xy}(l)$ is cross-covariance, σ^2_{EPO} and σ^2_{Hgb} are variances of resampled input and output sequence respectively, and l is lag between EPO and Hgb.

2.4.2 Output lables

To identify unacceptable model or acceptable model manually, the following sets of rules were used:

- If the magnitude of many hemoglobin 8-step prediction residuals are larger than +/-4, then a model is classified as an un-acceptable model
- If only a few of hemoglobin 8-step prediction residuals lie outside [-3, 3], then the model is classified as an acceptable model
- If there is not enough process excitation for input, classify as an un-acceptable model
- If model trend captures overall dynamics of actual data, classify as an acceptable model
- For an acceptable model, b parameters should increase and then decrease
- If the model has insignificant b parameters, classify as an un-acceptable model
- For an acceptable model, the sum of b parameters should be a positive number

These sets of rules gave 35 number of patients that resulted in unacceptable models from 145 patients in the first data set. Same sets of rules were used on the second data set. Table 2.1 summarizes output variable classification. Un-Acceptable models are referred to as 1 and acceptable models as 2.

Table 2.1: Number of acceptable and un-acceptable models for classification

Data Set Use	Data Set Number	Un-Acceptable(1)	Acceptable(2)
Test set	Data Set1 (145 patients)	35	110
Training & Validation set	Data Set2 (209 patients)	58	151

2.4.3 Data pre-processing

Table 2.2 summarizes range for different parameters in training set considered for model performance assessment. It can be observed that range for 16 parameters varies between 10^{-3} to 10^{4}. In order to reduce the effect of parameter magnitude different scaling methods were explored. Mean centering, auto-scaling, and range scaling was performed on the given data sets (Varmuza and Filzmoser, 2016).

- **Mean centering**: Mean centered variables were obtained by subtracting (arithmetic) mean of that parameter from all its values.

$$x_{ij}^{mc} = x_{ij} - \bar{x}(j) \tag{2.8}$$

where x_{ij}^{mc} denotes mean-centered variable, x_{ij} is original parameter value, $\bar{x}(j)$ is the mean value of that parameters, and i and j denotes patient number and variable number respectively. Figure 2.2 compares average mean centering value of patients on all 16 variables with raw data.

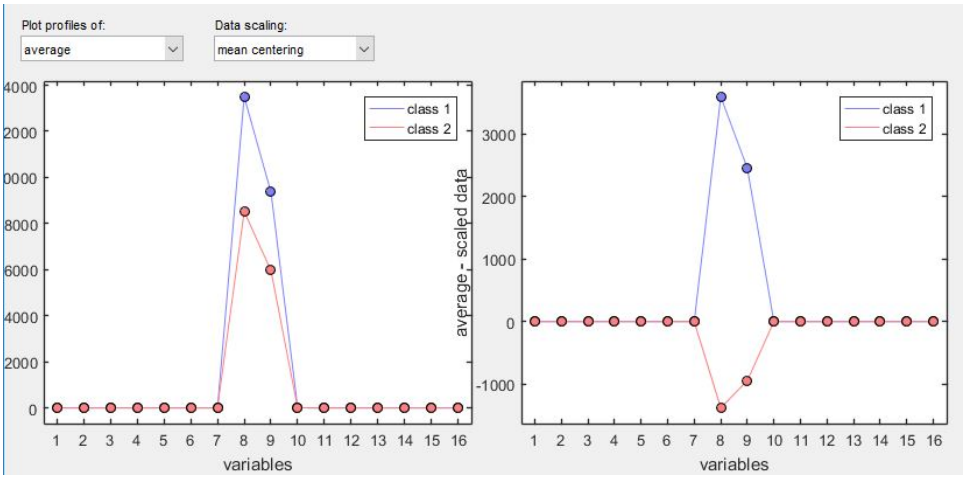


Figure 2.2: Raw Data vs Mean Centering

- **Auto scaling**: Auto scaled variables were obtained by subtracting (arithmetic) mean of that parameter from all its values and dividing with standard deviation.

$$x_{ij}^{as} = \frac{x_{ij} - \bar{x}(j)}{s_j} \tag{2.9}$$

where x_{ij}^{as} denotes auto-scaled variables and s_j denotes the standard deviation of particular parameter. Figure 2.3 compares average autoscaled value of patients on all 16 variables with raw data. All variables have been scaled from -1 to 1 magnitude.

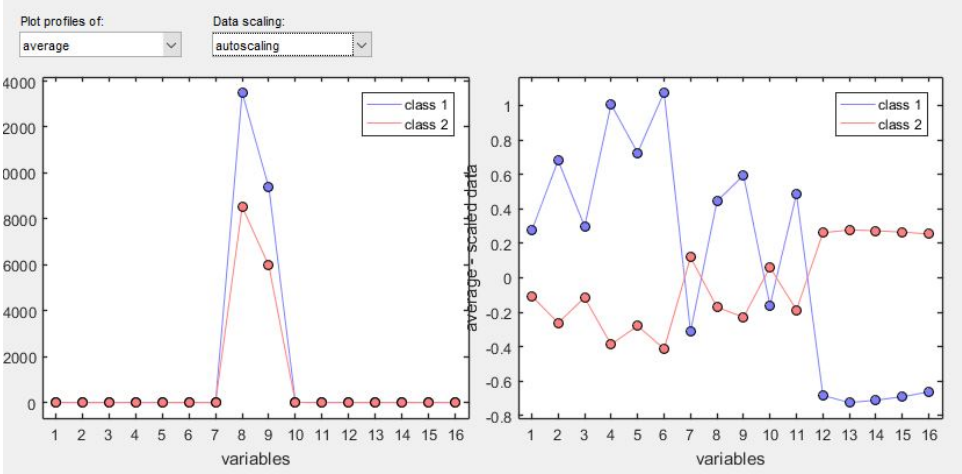


Figure 2.3: Raw Data vs Autoscaling

- **Range scaling**: Range scaled variables were obtained by subtracting minimum value of that parameter from all its values and dividing with range of that paramenter.

$$x_{ij}^{rs} = \frac{x_{ij} - x_j^{min}}{x_j^{max} - x_j^{min}} \tag{2.10}$$

where x_{ij}^{rs} denotes range-scaled variable, x_j^{max} and x_j^{min} are the maximum and minimum value of that parameter value respectively. Figure 2.4 compares average rangescaled value of patients on all 16 variables with raw data. All variables have been scaled from 0 to 1 magnitude. Range scaling method is explored to develop a linear model where all variables have same sign and comparable range. From the linear model, the sign and magnitude of the variable coefficient can signify type of effect it has on acceptable and unacceptable models. Final model is presented in section 2.5.3.

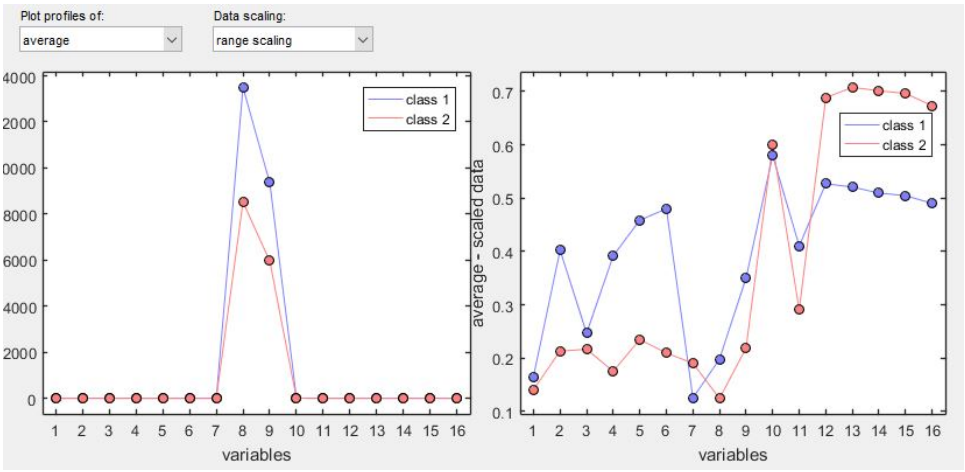


Figure 2.4: Raw Data vs Range-scaling

Table 2.2: Typical range of classification parameters

Parameter	Mean value	Maximum value	Minimum value
Mean hemoglobin 1-step prediction error	0.004928128	0.754402476	-0.124849983
Standard error of hemoglobin 1-step prediction errors	0.029762618	0.078696285	0.012047212
Mean hemoglobin 8-step prediction error	0.016819267	2.328439753	-0.654064857
Standard error of hemoglobin 8-step prediction errors	0.077186815	0.236298302	0.028154658
Hemoglobin 1-step standard deviation	0.345067575	0.805860099	0.150469632
Hemoglobin 8-step standard deviation	0.89680054	2.26649369	0.35165157
Sum of b parameters	0.200381319	0.656762826	0.1052625
Raw data input mean	9902.373837	67744.3609	96.15384615
Raw data input standard deviation	6912.050218	26187.45193	295.7515556
Raw data output mean	10.45171206	12.85642144	6.925118062
Raw data output standard deviation	1.004450741	2.203163287	0.429544064
CCF-1	0.017629412	0.387128672	-0.649816969
CCF-2	0.13032664	0.505736702	-0.585552352
CCF-3	0.218403739	0.627986624	-0.535734877
CCF-4	0.282901634	0.713663735	-0.493193028
CCF-5	0.322943242	0.787268265	-0.443107178

2.5 Classification techniques

Different classification methods were used for C-ARX model assessment to classify acceptable models (2) and un-acceptable models (1) mentioned in section 2.4.2. This section analyzes training as well as the test set on each of the following classification technique. Classification techniques handle class labels (supervised techniques) and define a mathematical model between descriptive variables and qualitative output variable (Ballabio and Consonni, 2013). In this section, exploratory analysis is presented, followed by supervised techniques to improve classification accuracy.

2.5.1 Exploratory analysis with principal component analysis (PCA)

PCA is a customary multivariate statistical method normally used to decrease the number of predictive variables and solve multi-colinearity problems. Principal component analysis finds un-correlated directions that can be utilized to summarize the original data without loosing too much information (Maitra and Yan, 2008).

In current patient data sets 1-step mean, 8-step mean communicate the similar type of information for model residuals. Similarly, 1-step standard error, 8-step standard error are effectively saying the same thing and capture variability in hemoglobin. Due to the existence of correlation PCA can be helpful to convert data to un-correlated variables. It might be suitable to transform the original set of variables to uncorrelated axes called PCs. These new axes (variables) will be a linear combination of original variables and are determined in order of importance such that first principal component captures maximum variation in the original dataset. Also, with PCA linear dimension reduction is possible which identifies orthogonal directions for maximum variational directions (Suryanarayana and Mistry, 2016).

Equations (2.11) to (2.15) (Jauregui, 2012) describe logical principle behind PCA. If matirx A contains n samples and m variables (Equation (2.11)), the mean of all variables can be stored in a single vector $\vec{\mu}$ (Equation (2.12)). After centering the data matrix A can be transformed to B (Equation (2.13)) so that each variables have a mean as zero. Equation (2.14) defines the *covariance matrix* S (m x m symmetric

matrix) which is orthogonally diagnolized in Equation (2.15).

$$A = [\vec{x_1}, \vec{x_2}, \vec{x_3}, ..., \vec{x_n}] \tag{2.11}$$

$$\vec{\mu} = \frac{[\vec{x_1} + \vec{x_2} + ... + \vec{x_n}]}{n} \tag{2.12}$$

$$B = [\vec{x_1} - \vec{\mu}|....|\vec{x_n} - \vec{\mu}|] \tag{2.13}$$

$$S = \frac{1}{n-1} BB^T \tag{2.14}$$

$$S\vec{v_i} = \lambda_i \vec{v_i} \tag{2.15}$$

where $\lambda_1 \geq \lambda_2 \geq ...\lambda_m \geq 0$ be the eigenvalues of S (in decreasing order) with corresponding orthonormal eigenvectors $\vec{u_1}, \vec{u_2}, , ...\vec{u_m}$. These eigenvectors are known as *prinicpal components* of the data set and sum of eigenvalues ($\lambda_1 + \lambda_2 + ... + \lambda_m = T$) describes *total variance* of the data set A. Jauregui, 2012 summarized following sets of interpretation that can be drawn from PCA:

- The first direction $\vec{u_1}$ (the first principal component/direction) explains maximum variance (λ_1) for the data set and contributes $\frac{\lambda_1}{T}$ variance. Similary, $\vec{u_2}$ (the second principal component/direction) contributes $\frac{\lambda_2}{T}$ variance.

- The eigenvectors $\vec{u_1}, \vec{u_2}, , ...\vec{u_m}$ are new uncorrelated directions corresponds to new variables for the data set.

- If initial λ_i's are much bigger than all other eignevalues, dimension reduction is possible.

Analysis on dataset 1

PCA on data set 1 with original parameters gave some separation on new uncorrelated PC directions. Eigenvalues for PCA on this data set are:

[35.25, 25.62, 13.43, 8.06, 7.61, 5.95, 2.38, 0.77, 0.49, 0.25, 0.10, 0.02, 0.005]

It can be seen that the first components contribute 35% of the variance. This can be observed in Figure 2.5, where some separation can be seen on PC1. Because of only small contribution toward total variance, more than one PC direction need to be considered.

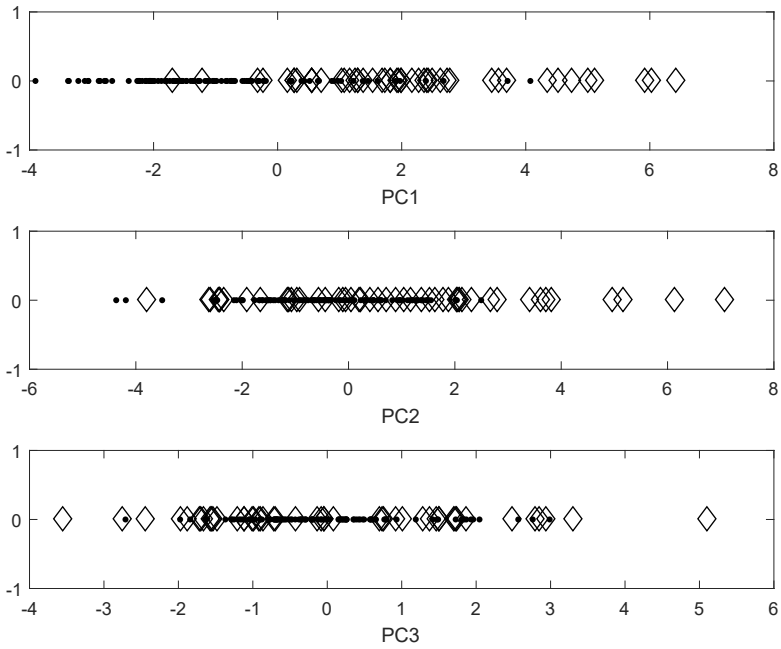


Figure 2.5: Raw Separation in 1-dimension. (.) unacceptable, ◊ acceptable model

Increasing more PC components improves separation of data and some common features can be observed. The first two components contributes 61% of the variance, whereas adding the third PC cumulates to 74% of the variance. This can be observed in Figures 2.6 and 2.7, where more separation can be seen between classes.

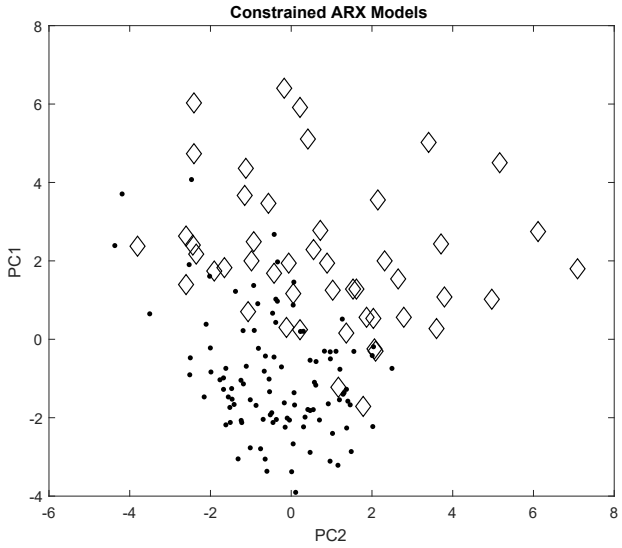


Figure 2.6: Raw Separation in 2-dimension. (.) unacceptable, ◊ acceptable model

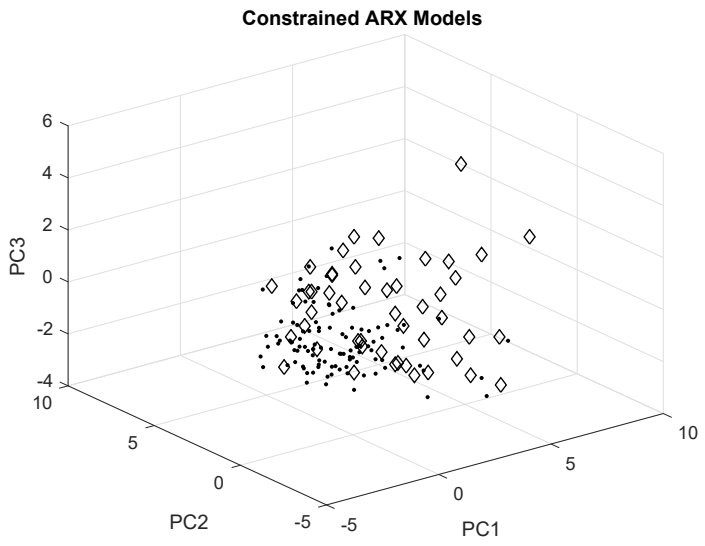


Figure 2.7: Raw Separation in 3-dimension. (.) unacceptable, ♢ acceptable model

Limitation with PCA on clinical data set

Although PCA helped to transform correlated parameters to orthogonal PC components, there was no mention of class labels in Equations (2.11) to (2.15) (Jauregui, 2012). For this reason, PCA is also called as unsupervised technique (Jolliffe, 1986). From Figures 2.5 to 2.7 it can be seen that the two classes have some overlap for model classification. This can be assigned to the limitation of PCA as it was not trained with class labels. Therefore, supervised techniques in next sections (PLS-DA and LDA) were expected to increase this separation.

2.5.2 Partial least squares-discriminant analysis (PLS-DA)

PLS-DA is a linear classification technique to use the power of discriminant analysis along with least square regression. In PLS-DA original data variability is modeled with Latent Variable (LVs), linear combinations of original parameters, which permits graphical perception and comprehension of the distinctive patterns and relations with scores and loadings. Linear combination of original variables gives coefficient of loadings which decide the LVs. Loadings represent the effect of each variable on LV. Scores are the new coordinates of samples in LV space (Ballabio and Consonni, 2013).

Equations (2.16) to (2.18) (Wold *et al.*, 2001) describes principle behind PLS-DA. If X denotes measurement matrix with n samples and m variables and Y denotes

class labels (n x 1), then X-scores which are latent variables (LVs) are calculated as linear combination of original variables. W contains weights that combines with original variables for LVs.

$$\begin{aligned} T &= XW^* \\ X &= TP' + E \\ Y &= UC' + G \end{aligned} \tag{2.16}$$

T and U are scores on X and Y. P and C corresponds to loadings. X-scores are good predictors of Y so that Y-residuals (F) are "small", i.e.:

$$Y = TC' + F \tag{2.17}$$

E and F represent residuals on X and Y respectively. Equations (2.16) to (2.17) can be re-written as

$$\begin{aligned} Y &= XW^*C' + F = XB + F \\ B &= W^*C' \end{aligned} \tag{2.18}$$

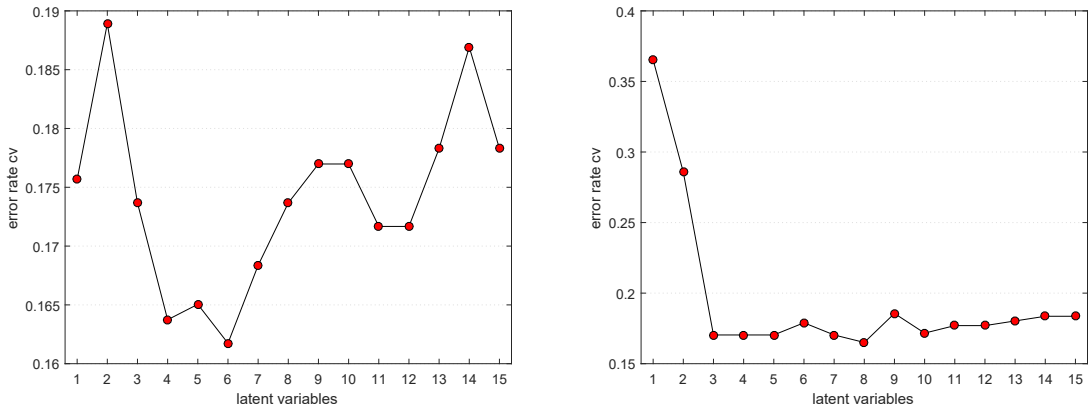
B represents "PLS-regression coefficients". Once Y is calculated, a threshold can be defined to minimize number of false positives and false negatives. This threshold can be moved to either side based on requirement of process.

Selection of latent variables from training set

Data set 2 was used as training set. Optimum number of LVs were calculated using cross validation (CV) procedure. For patient data set total number of samples (209) in training set 10 CV groups were used. Figures 2.8a and 2.8b show plot of CV error to find optimal number of LVs. Figures 2.8a and 2.8b correspond to mean centering and autoscaling methods respectively. Error rate cv is calculated as Equation (2.19).

$$\text{error rate cv} = \frac{\text{No. of misclassifications}}{N} \tag{2.19}$$

In Figure 2.8a minimum error rate is 0.1617 with six LVs. Similarly, with mean-centering error rate decreases from 0.36 to 0.17 with 3 LVs. Increasing LVs further do not reduce the error rate. Since autoscaling scaled variables to the same range of magnitude, autoscaling with 6 LVs was used for developing regression model.



(a) autoscaling **(b)** mean-centering

Figure 2.8: Optimal number of LVs with 10 cross validation groups

PLS-DA model calculation and analysis

Once the optimal number of the component has been decided, PLS-DA regression model can be calculated selecting 6 LVs with autoscaling and 10 Cross-validation group. Table 2.3 shows the confusion matrices obtained from training.

Table 2.3: Confusion Matrix for training set

Actual/Predicted	Un-Acceptable(P)	Acceptable(N)
Un-Acceptable(P)	47 (TP)	11 (FN)
Acceptable(N)	17 (FP)	134 (TN)

where TP (True Positive) is the number of unacceptable models predicted correctly and TN (True Negative) is the number of acceptable models predicted correctly. FN (False Negative) is the number of unacceptable models that are predicted as acceptable and FP (False Postive) is the number of acceptable models that are predicted as unacceptable (Ballabio and Consonni, 2013). *Senstivity* (S_n) is calculated as TP/(TP + FN). Senstivity close to 1 represents model ability to recognize unacceptable model. In contrast, *specificity* (S_p) is calculated as TN/(FP + TN), higher the specificity better the model is to recognize acceptable model. The ratio of (TP + TN) to (Total Models) is ratio of correct assignments, which is called accuracy (Ballabio and Consonni, 2013). Table 2.4 summarize S_n, S_p, and accuracy for training set.

Table 2.4: Error rates in training

Accuracy	Senstivity	Specificity
0.87	0.81	0.89

For the patient data set, it is required to have high sensitivity because all un-acceptable models should be classified as un-acceptable. This is due to the risk of using a bad model for a particular patient. For acceptable model, in case it is classified as unacceptable model, physician will has to wait for more data and keep using the previous model until the new acceptable model is available.

Besides accuracy, sensitivity, and specificity, ROC (Receiver Operating Characteristics) graphical tools can be used to assess classification quality. ROC curves are shown in Figure 2.9 which are drawn as plot between sensitivity vs 1-specificity. For perfect separation, left-hand top corner would be reached with ROC area (Ballabio and Consonni, 2013). As expected ROC area under the curve (AUC) is not equal to 1 but classification results are still acceptable with AUC of 0.95. The curve on the right-hand side shows Sensitivity and Specificity with changing threshold. The point where both lines cut each other is the point where the number of FP and FN are minimized. From the figure, the optimum threshold for PLS-DA classification is 0.04. This means that y_i^{cal} greater than 0.04 will be classified as an acceptable model.

The last important statistics for PLS-DA is the amount of variance explained by LVs. From Figure 2.10 it can be seen that 6 LVs corresponds to 92% of the variance on measurement data, X, which is a good approximation of original data. The explained variance with 6 LVs on Y (class data) is 20% and not much addition to variance by following LVs. This is quite common with PLS-DA where acceptable classification model can be found without explaining much variance (Ballabio and Consonni, 2013).

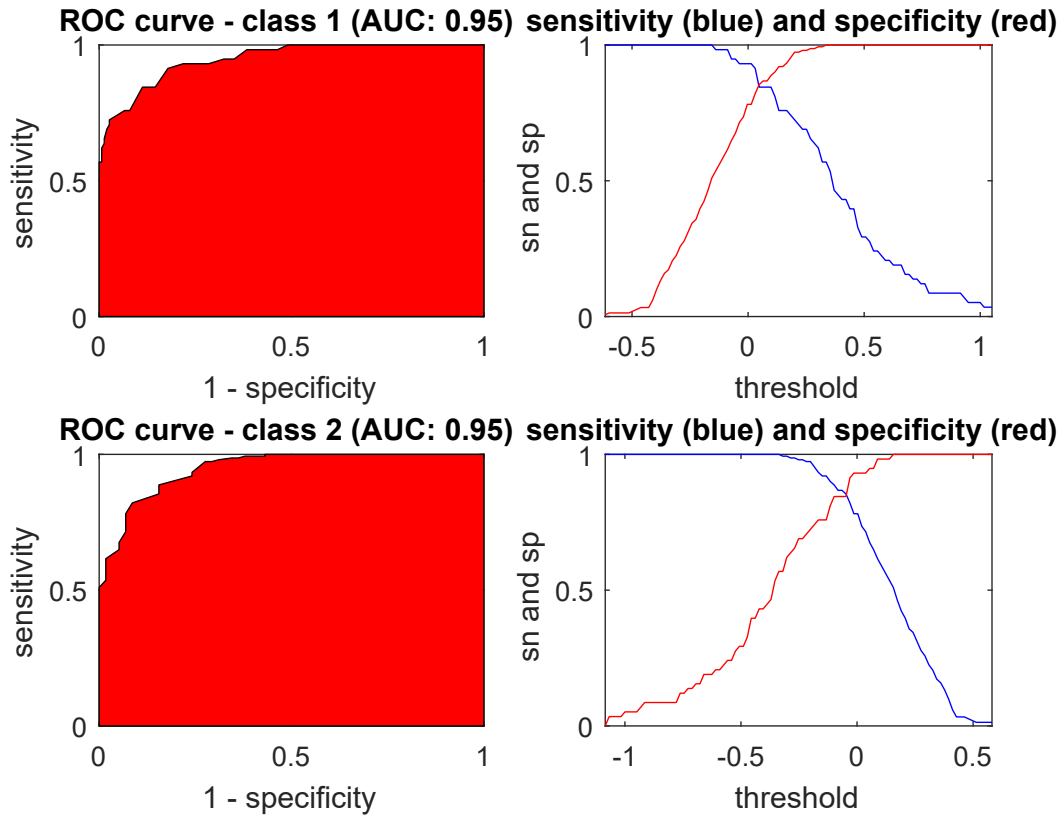


Figure 2.9: ROC curves(left) and plots of S_n (blue) and S_p (red) values as threshold is changing (right) for class-1:unacceptable (upper), class-2:acceptable models (lower)

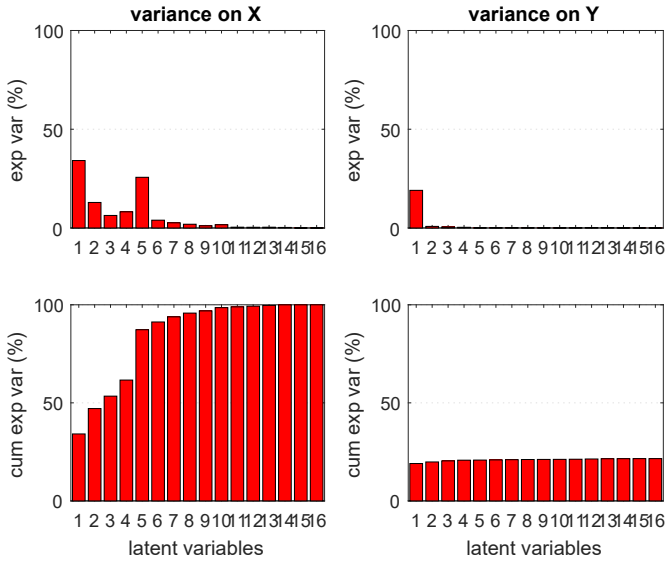


Figure 2.10: Amount of variance explained by LVs

Scores plot on LV1 and LV2 are shown in Figure 2.11. Total variance explained with these two components is 47%.

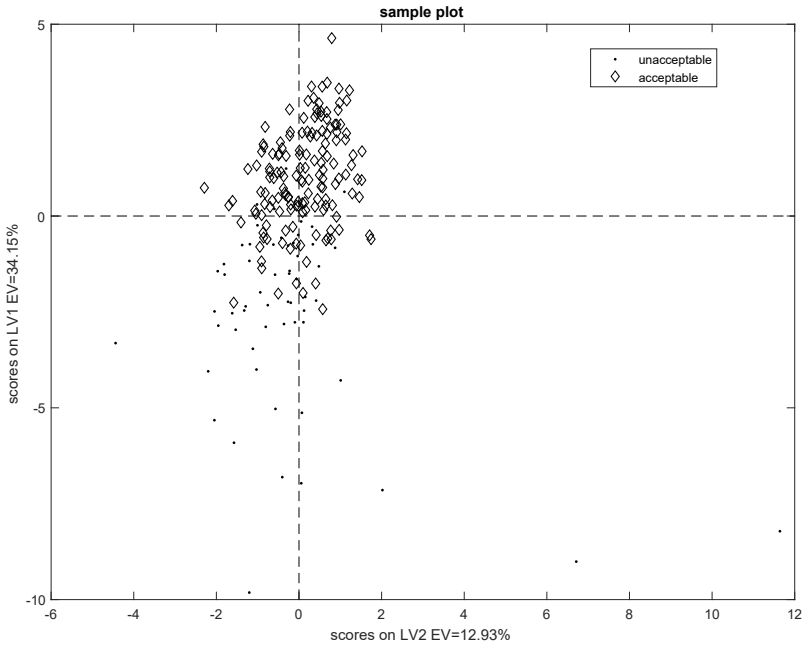
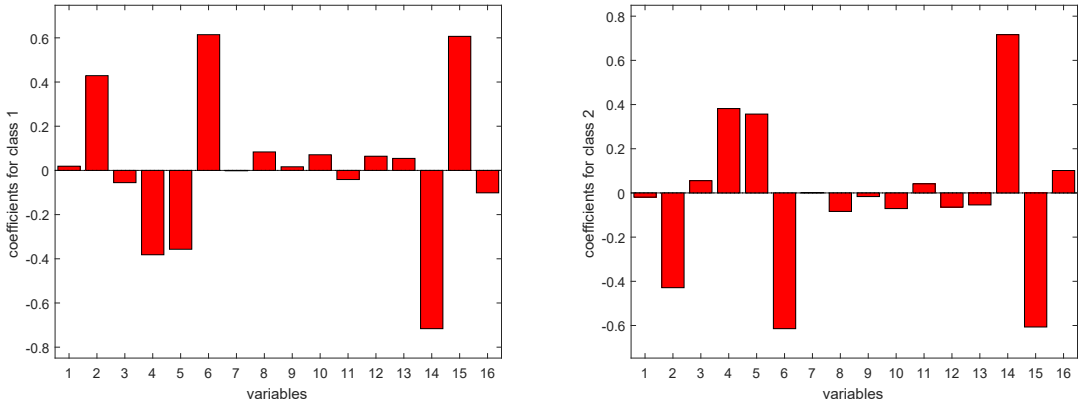


Figure 2.11: Score plot with first two LVs retaining 47% of variance

Figures 2.12a and 2.12b show regression coefficients for class 1 and class 2 respectively. Since there are no un-assigned sample, regression coefficents are simply negative of one another.



(a) class1:un-acceptable models **(b)** class2:acceptable models

Figure 2.12: Regression coefficents from PLS-DA for both classes

Test set performance analysis

Dataset 1 is used as the test set and was kept separate from datset 2. Dataset 1 contains 145 sample patients with 35 un-acceptable models and 110 acceptable models. Since these samples were not used for PLS-DA model calculation, they will be used now to assess predictive ability of the built model. Tables 2.5 and 2.6 summarize performance on test set. Accuracy, sensitivity, and specificity are comparable with training set results. Therefore, PLS-DA model developed can be considered as reliable and stable.

Table 2.5: Confusion Matrix for test set

Actual/Predicted	Un-Acceptable(P)	Acceptable(N)
Un-Acceptable(P)	28 (TP)	7 (FN)
Acceptable(N)	16 (FP)	94 (TN)

Table 2.6: prediction error rates

Accuracy	Senstivity	Specificity
0.84	0.80	0.85

Limitation with PLS-DA results on clinical data set

Although PLS-DA results gave 84% accuracy for class-separation, higher performance was expected. Since we would not like to use bad model for actual patients, greater than 90% accuracy was targeted for higher confidence to use classification techniques. In text section, another technique: combination of Principal Component Analysis along with Linear Discriminant Analysis is explored to achieve high accuracy performance.

2.5.3 Principal component analysis - Linear discriminant analysis (PCA-LDA)

Similar to PCA, Fisher Linear Discriminant Analysis (also called as Linear Discriminant Analysis (LDA)) is another tool used for classification. LDA is closely related to PCA and both methods are linear. In PCA, the mean square error between original data and projected data on new transformations is minimized (Li and

Wang, 2017). Also, PCA is a un-supervised technique which does not take into account class labels (Jolliffe, 1986). But for LDA, matrix transformation is done by maximizing the ratio of between-class variance to within-class variance. This goal increases the separation between classes at the same time reducing within class variation (Li and Wang, 2017).

LDA model theory

Li and Wang, 2017 explains LDA as follows from Equations (2.20) to (2.26). If y is output variable and X is measurment matrix with m - dimensional samples:

$$X = [x^1, x^2, ...x^m] \tag{2.20}$$

and N_1 of total samples belong to class-1 and N_2 belong to class-2. LDA seeks to obtain y by projecting X onto a line with slope θ slope as:

$$y = \theta^T X \tag{2.21}$$

Mean vector of two classes in X-space (μ) will be as follow:

$$\mu_k = \frac{1}{N_k} \sum_{i \in C_k} x^i \qquad \text{where } k = 1,2 \tag{2.22}$$

and in y-space ($\hat{\mu_k}$):

$$\hat{\mu_k} = \frac{1}{N_k} \sum_{i \in C_k} y^{(i)} = \frac{1}{N_k} \sum_{i \in C_k} \theta^T x^i = \theta^T \mu_k \quad \text{where } k = 1,2 \tag{2.23}$$

LDA maximizes **variance between classes** as distance between class means (μ_k) as:

$$\hat{\mu_2} - \hat{\mu_1} = \theta^T (\mu_2 - \mu_1) \tag{2.24}$$

and minimizes **within class-variance** (s_k) for each class as:

$$\hat{s_k^2} = \sum_{i \in C_k} (y^{(i)} - \hat{\mu_k})^2 \qquad \text{where } k = 1,2 \tag{2.25}$$

Then, according to LDA definition - maximizing the ratio of between-class variance to within-class variance, objective function ($J(\theta)$) is defined as follow:

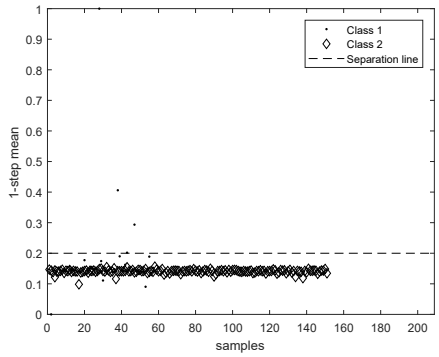
$$J(\theta) = \frac{(\hat{\mu_2} - \hat{\mu_1})^2}{(\hat{s_1^2} + \hat{s_2^2})} \tag{2.26}$$

PCA-LDA algorithm

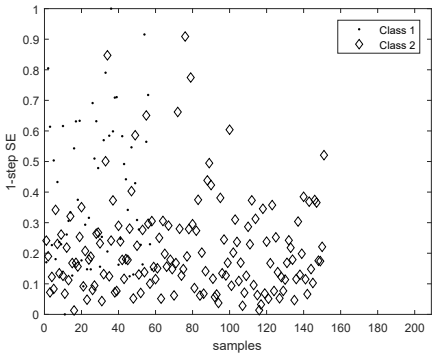
PCA followed by LDA is often used technique when discriminant power lies in first few orthogonal components of PCA (Næs and Mevik, 2001). First, PCA is performed on the whole data set and then LDA is used on new orthogonal directions given by PCA. Goal for the patient data set is to increase class separation so that correct model should be used for real patients. This technique often result into better classification result.

Preliminary data analysis to detect un-acceptable model

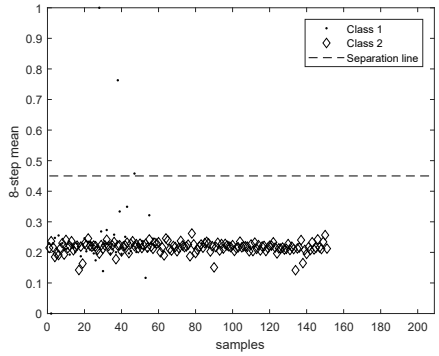
Figures 2.13 and 2.14 contains a plot of all sixteen range scaled variables on the training set. The black line roughly separates outliers, patients which are generally unacceptable. These cut-off on the training set will be used to detect unacceptable models.



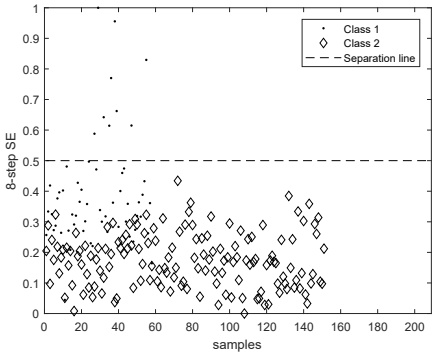
(a) 1-step mean



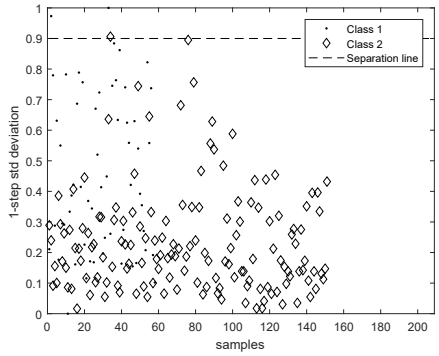
(b) 1-step standard error



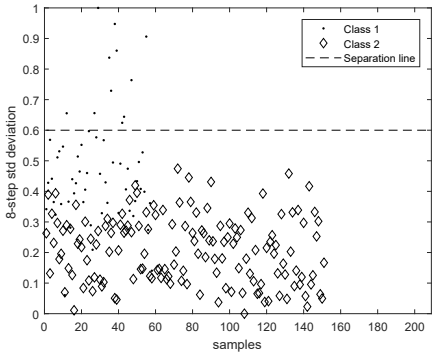
(c) 8-step mean



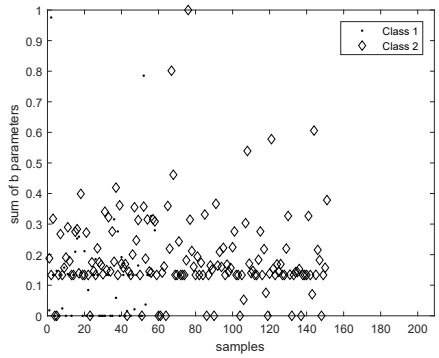
(d) 8-step standard error



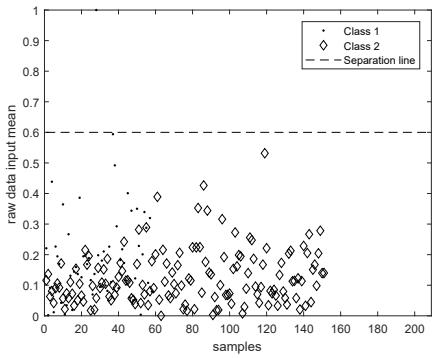
(e) 1-step standard deviation



(f) 8-step standard deviation

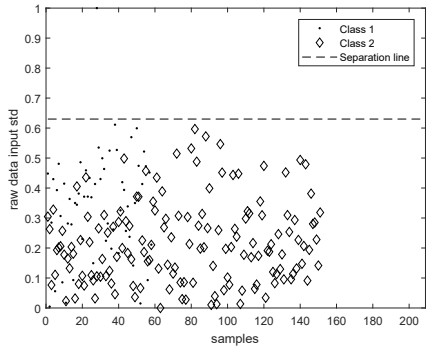


(g) sum of b parameters

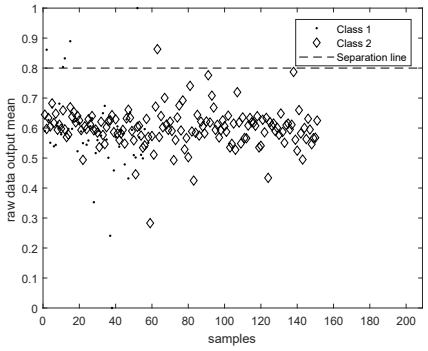


(h) raw data input mean

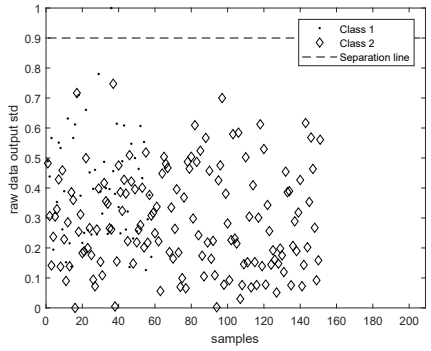
Figure 2.13: preliminary data separation based on magnitude of variables (1 to 8)



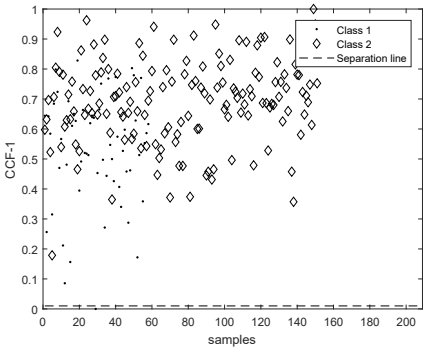
(a) raw data input std deviation



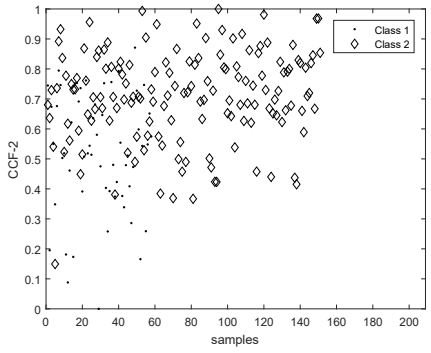
(b) raw data output mean



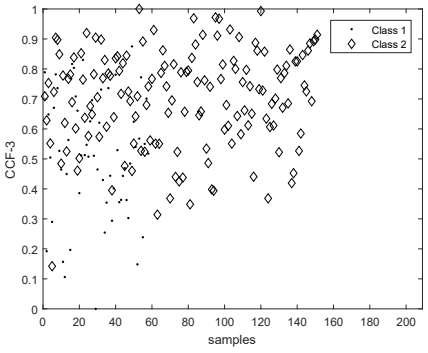
(c) raw data output std deviation



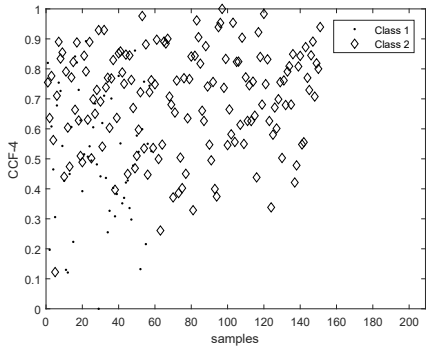
(d) CCF 1



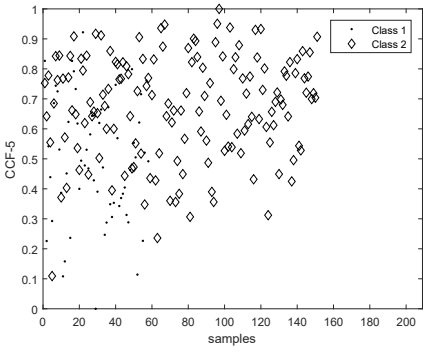
(e) CCF 2



(f) CCF 3



(g) CCF 4



(h) CCF 5

Figure 2.14: Data separation based on magnitude of variables (9 to 16). Variables are range scaled. Black line is rough separation above which models are generally un-acceptable

Selection of PC components from training set

The optimal number of LVs were calculated and presented in Figure 2.15. Similar to PLS-DA cross-validation was performed on the training set using 10 CV groups. It can be observed that error rate is minimum with 11 PC components. After 11 components error rate increases. Therefore, 11 PC components were used for subsequent LDA analysis.

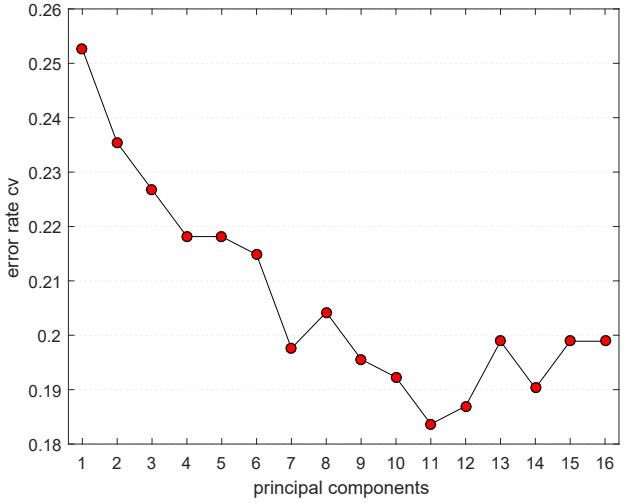


Figure 2.15: Optimal Number of PC components

From Figure 2.16 it can be seen that 11 PC components corresponds to more than 99% of variance, which justifies most of the information retained from original data.

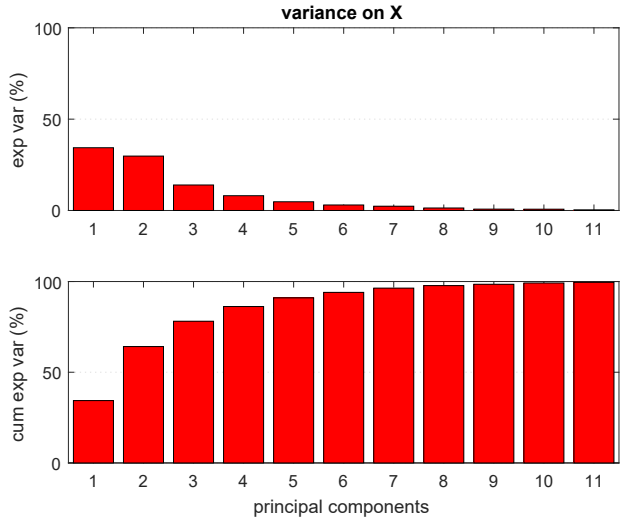


Figure 2.16: Amount of variance explained by selected PC components

PCA-LDA calculation for binary classification

Once the optimal number of PC components have been decided, LDA can be performed using the orthogonal PC components. Once PCA is done, let L be number new PC directions. Scores on new PC directions can be found as:

$$T = XL \tag{2.27}$$

where X is original scaled data matrix, L: loadings from PCA. Let μ_1 and μ_2 be mean vector for two classes, T_1 and T_2 are scores for class-1 and class-2. S_1 and S_2 be covariance matrices of T_1 and T_2, then with-in class scatter matrix, S_w:

$$S_w = \frac{S_1 N_1 + S_2 N_2}{N} \tag{2.28}$$

where N is total number of samples, N_1 of which belong to class-1 and N_2 to class-2. LDA separates binary classes by projecting samples onto line (Xiong and Cherkassky, 2005):

$$\begin{aligned} y &= a_0 + a_1' T' \\ \text{where, } a_o &= log(p_1/p_2) - \frac{1}{2}(\mu_1 + \mu_2)' S_w^{-1}(\mu_1 + \mu_2) \\ a_1 &= S_w^{-1}(\mu_1 + \mu_2) \end{aligned} \tag{2.29}$$

p_1 and p_2 are empirical frequencies of each class in training set

$$p_i = \frac{N_i}{N} \qquad \text{where, } i = 1,2 \tag{2.30}$$

with back substitution, y can be founds as a function of original parameters in Equation (2.31)

$$\begin{aligned} y &= a_0 + bX' \\ \text{where, } b &= a_1' L' \end{aligned} \tag{2.31}$$

From the above model sample will be assigned to class 1 if $y > 0$ and to class 2 otherwise. The final PCA-LDA model coefficients are shown in Figure 2.17. It can be seen that high *8-step error standard deviation* contributes towards un-acceptable model, followed by *raw-output mean*. All cross-correlation functions found to have negative coefficents except *CCF5*.

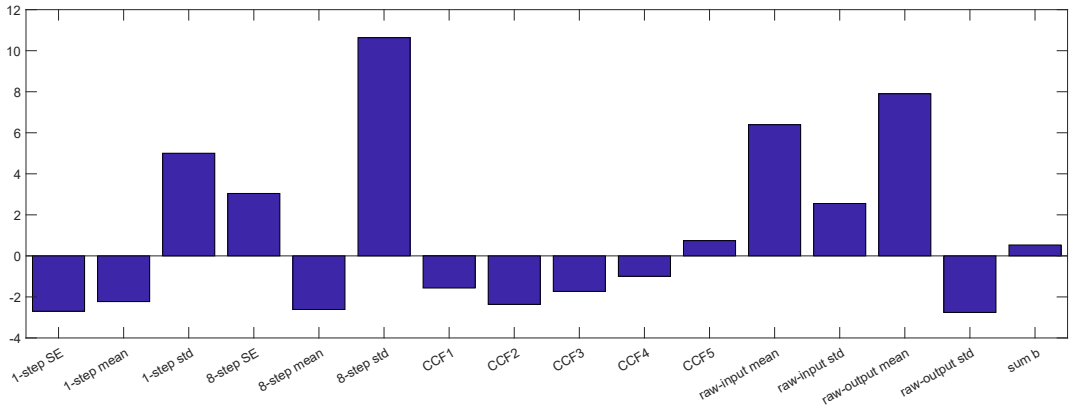


Figure 2.17: PCA-LDA model coefficients, range scaled X

Confusion matrix for the above PCA-LDA model along with preliminary data analysis from Figures 2.13 and 2.14 is summarized in Table 2.7.

Table 2.7: PCA-LDA Confusion Matrix for training set

Actual/Predicted	Un-Acceptable(P)	Acceptable(N)
Un-Acceptable(P)	41 (TP)	17 (FN)
Acceptable(N)	7 (FP)	144 (TN)

The overall performance of training set in terms of accuracy, senstivity, and specificity is presented in Table 2.8. Specificity has increased by 10% as compared to PLS-DA model presented in Table 2.4, whereas sensitivity to recognize un-acceptable model has to decreased to 0.71 against 0.80 with PLS-DA. But overall performance accuracy has increased from 84% to 89%.

Table 2.8: PCA-LDA error rates in training

Accuracy	Senstivity	Specificity
0.89	0.71	0.95

Test set performance analysis

Tables 2.9 and 2.10 summarizes performance on test set (Data set 2). Accuracy, senstivity, and specificity is comparable and better than training set results. Therefore, PCA-LDA model developed can be considered as reliable and stable.

Table 2.9: PCA-LDA Confusion Matrix for test set

Actual/Predicted	Un-Acceptable(P)	Acceptable(N)
Un-Acceptable(P)	31 (TP)	4 (FN)
Acceptable(N)	6 (FP)	104 (TN)

Table 2.10: PCA-LDA prediction error rates

Accuracy	Senstivity	Specificity
0.93	0.89	0.95

2.6 Conclusions

In this chapter, three different classification techniques were explored. These classification techniques were used for model performance assessment for constrained ARX model. Performance assessment was done on 209 clinical patient data set and the prediction was done on another 145 patients. Exploratory analysis with Principal Component Analysis (PCA) gave some separation between the two classes, but not very distinctive separation because of its limitation, being un-supervised technique. Partial Least Square-Discriminant Analysis (PLS-DA) gave better performance than PCA with six LVs explaining 92% of the variance in the original data matrix. Overall accuracy of assigning the new patient to the correct class with PLS-DA is 84%. PLS-DA prediction performance for the acceptable model (specificity) and un-acceptable model (sensitivity) are 80% and 85% respectively. The final technique explored was PCA-LDA. This combination of principal component analysis and linear discriminant analysis gave better overall performance than PLS-DA and PCA. Overall prediction accuracy got increased to 93%, specificity to 95%, but sensitivity to 89%.

Taken altogether, in order to predict whether C-ARX model will result in an acceptable model or unacceptable model, it is recommended to use PCA-LDA technique because of its simplicity and high prediction accuracy.

Chapter 3

System Identification with Iron Level and WBC

3.1 Introduction

This chapter is focused on system identification of hemoglobin models with additional parameters to improve constrained ARX model developed by J. Ren *et al.*, 2017. This includes adding additional parameters such as Iron Saturation (TSAT) (also known as transferrin saturation), White-Blood-Cells (WBC), and Ferritin level in the constrained optimization problem. These measurements were available for clinical data set 1 (167 patients). Improved model identification was done in three parts: first, by introducing TSAT in the model; second, by optimizing with TSAT along with WBC; and third by including serum ferritin instead of TSAT. It will be shown that simultaneous optimization, or sequential optimization with TSAT, or original constrained ARX model can be used under different conditions to improve model performance.

From Figure 1.3 it can be observed that physicians monitor TSAT and ferritin levels before EPO dosage initiation. This explains the dependancy of hemoglobin level on TSAT and ferritin. In the following sub-sections, literature review is presented for considering TSAT, WBC, ferritin, and iron dosage in hemoglobin modeling.

3.1.1 Iron saturation and ferritin serum level

Iron level in CKD patients is very important. Due to frequent blood testing, hemodialysis and other tests, CKD patients suffer from the iron loss that can range

from 1.5 to 3 g/yr (Kalantar-Zadeh *et al.*, 2006). Erythropoietin treatment without sufficient iron level in the body does not raise hemoglobin level. On the other hand, iron supplementation without EPO dosage is unsuccessful because erythropoietin acts as a contractual worker to make Hgb level in blood veins (Kalantar-Zadeh *et al.*, 2006).

Iron level stored in the body is reflected by TSAT and serum ferritin. TSAT is calcualted from serum level and Total Iron Binding Capacity (TIBC) as shown in Equation (3.1) (Group *et al.*, 2001).

$$\text{TSAT} = \frac{\text{serum iron} \times 100}{\text{TIBC}} \tag{3.1}$$

Ferritin or serum ferritin is also reflector of iron storage in human body. Ferritin is a cell in blood (called as blood cell protein) containing iron in it. It is mainly present in bone marrow, liver, spleen, etc (Walters *et al.*, 1973). A comparision is presented between TSAT and serum ferritin advantages and limitations by (Kalantar-Zadeh *et al.*, 2006) in Table 3.1.

Table 3.1: Comparision between TSAT and serum ferritin

Iron Store label	Advantages	Limitations
TSAT	More sensitive and reliable than ferritin	Denominator (TIBC) can be low due to other reasons such as malnutirtion and/or inflammation
Serum ferritin	Low level highly signify iron deficiency	Moderately high values can be because of non-iron-related scenarios

National Kidney Foundation recommends a TSAT levels between 20 to 50% and serum ferritin level between 100 to 800 ng/mL for patients suffering from CKD (National Kidney Foundation Anemia Working Group *et al.,* 2001). They also found that increase in TSAT beyond 50% or serum ferritin level increase to 800 ng/mL, CKD patients do not respond to hemoglobin level with increase in EPO dosage. Also, TSAT and serum are recommended to be monitored every 3 to 6 months.

On the other hand, World Health Organisation (WHO) demonstrated that ferritin level greater than 200 ng/mL for men and 150 ng/mL for women are related to iron overdose (Organization *et al.*, 2011). This leads to deceptive information from fascinating TSAT and serum ferritin level. Uncertainty in this information concludes

that there is no specific range that than is generalized to all patients. The optimal range of TSAT and ferritin level can be different for different patients. For this reason, in subsequent sections, TSAT and ferritin were set as optimization variables to find the optimal range for a particular patient. A high iron level is associated with no effect on EPO, whereas a low level of iron is assumed to have less effective exogenous EPO level. Since both ferritin and TSAT reflects stored iron in the body (National Kidney Foundation Anemia Working Group *et al.,* 2001), only one was introduced in the optimization problem each time.

3.1.2 Exogenous iron dosing

CKD patients commonly suffer from iron deficiency, especially if they undergo hemodialysis. This can be due to different reasons such as loss of blood from frequent blood tests, blood loss during dialysis, blood loss in the intestine, etc. Also, epoetin dosing treatment increases erythropoiesis activity rate, which subsequently increases the demand for iron. This increased demand along with blood loss, aggravates the trouble of keeping up sufficient iron level (Ferritin 100 to 800 ng/mL and TSAT 20 to 50%) in the body (National Kidney Foundation Anemia Working Group *et al.,* 2001).

The National Kidney Foundation made recommendations for iron dosing considering following issues (National Kidney Foundation Anemia Working Group *et al.,* 2001):

1. The iron requirements and significance of keeping iron status at the satisfactory level.

2. Evaluation of iron level: sensitivity of TSAT and serum ferritin to detect deficiency of iron and/or iron overload.

3. An evaluation for the mode of iron treatment: oral versus IV iron

4. Administering IV iron and potential risk factors

5. Overload by iron dosing (National Kidney Foundation Anemia Working Group *et al.,* 2001).

The clinical data set used in subsequent sections had intermittent dosings of iron. Patients have 100 mg of IV iron every other day on an average over 2 weeks and then dosing was stopped and started again when TSAT or ferritin level falls below the recommended level. This resonates with guideline F. mentioned by National Kidney Foundation (National Kidney Foundation Anemia Working Group *et al.,* 2001), where IV iron is given if TSAT falls below 20% and/or ferritin below $100ng/mL$ at each hemodialysis for 8 to 10 doses. IV iron could be included in modeling for a multi-input-multi-output (MIMO) system where IV iron and exogenous EPO could be considered as input and hemoglobin and TSAT as outputs. This model was motivated from Gaweda *et al.*, 2014, which is based on iron cycle and erythropoietin cycle in body Figure 3.1 (Gaweda *et al.*, 2014).

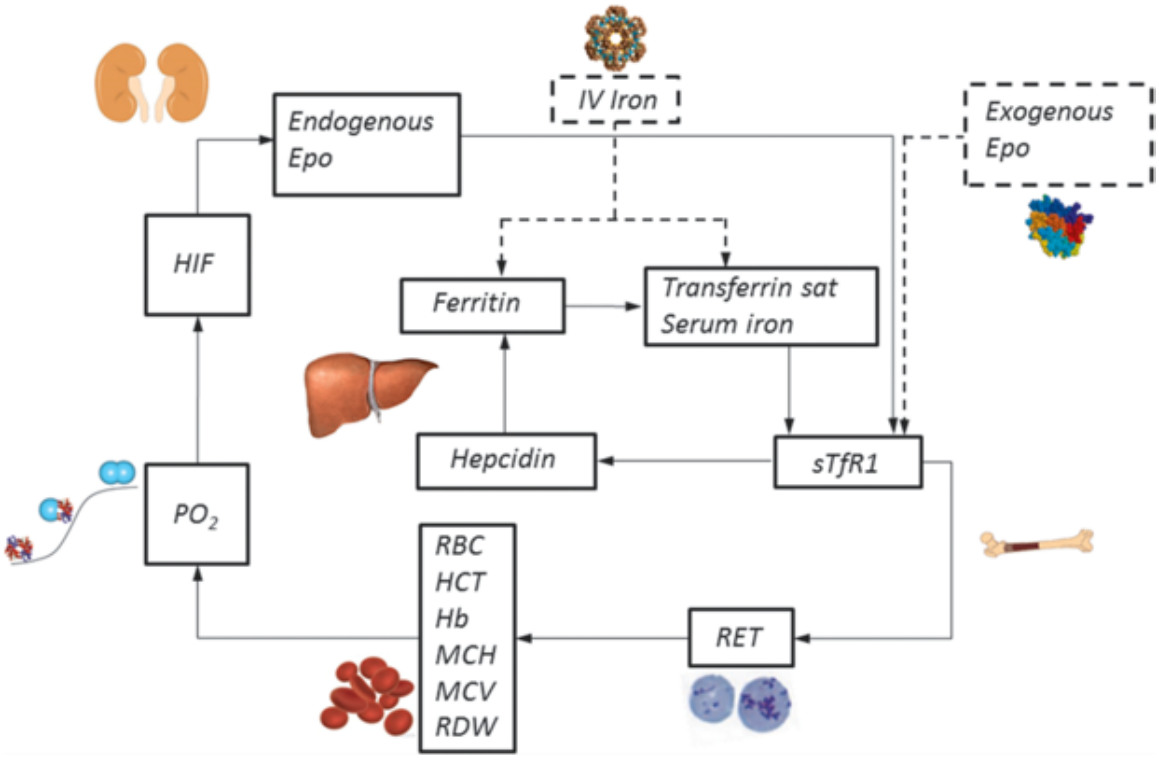


Figure 3.1: Mathematical models of ESA and/or IV iron dosing, taken from (Gaweda *et al.*, 2014)

3.1.3 White blood cell level (WBC)

When a patient undergoes Complete Blood Count (CBC), examination shows red blood cells count, hemoglobin, hematocrit, WBC, and platelets in the blood. Around

1% of total blood volume are WBCs. The role of WBCs is to fight against infection and demolish microbes and germs that infect by entering the body (George-Gay and Parker, 2003). WBCs are also called as leukocytes. There are total five different types of WBCs in body (George-Gay and Parker, 2003): Neutrophils, Basophils, Eosinophils, Lymphocytes, and Monocytes. With the help of these five types of cells, WBC drives human response system (George-Gay and Parker, 2003). Units of measurement for WBCs is K/μL and normal range is 4.5 to 10 K/μL (George-Gay and Parker, 2003).

Elevated WBC level

The elevated level of WBC count greater than 11 K/μL WBC count is called *leukocytosis*. If there is an acute infection, bone marrow starts increased production of WBCs in blood. Leukocytosis is the most common type of infection. Also, WBC count greater than 10,000 has been related to high mortality rates (George-Gay and Parker, 2003). Hence, in subsequent sections effect of elevation in WBC levels is negatively associated with hemoglobin and WBC is also selected as optimization variables. Opposite to iron level in body, high WBC level is associated with less effective EPO.

3.2 System identification to improve existing C-ARX model

In this section, C-ARX model developed by J. Ren *et al.*, 2017 will be used as a base model and measurements such as TSAT, ferritin, and WBC will be used to include and to improve model performance on hemoglobin. The frequency of measurement of these additional parameters are every 2 to 3 months.

For the purpose of simplicity and consistency, to compare modeling results with J. Ren *et al.*, 2017, it is expected that the week by week hemoglobin estimations are present. Ideally, clinical sampling of hemoglobin is four times in each month (A. Gaweda *et al.*, 2010). The raw clinical data used for the model update is same that was used by J. Ren *et al.*, 2017 and was re-sampled. The EPO dosing time also varied for different patients with general dosing frequency being weekly. From

Figures 3.2 and 3.3, it can be observed that TSAT, ferritin, and WBC have smaller measurement frequency than Hgb and EPO. For this reason, re-sampled values were used for all parameters in modeling, where Hgb, TSAT, ferritin, and WBC values were linearly interpolated from original data. Drug dosages (EPO and IV iron dosing) were re-sampled into a week by week fashion by taking the aggregate of the present day drug level and the past 6 days of measurements. Also, from Figures 3.2 and 3.3, it can be seen that ferritin and TSAT measurements are not available at the beginning of data as well as towards the end. For this reason, modeling was carried out from the first available measurement of TSAT and ferritin till the last available measurement, and rest of the data was excluded from modeling. This approach decreases the amount of data available for modeling but 80 to 100 weeks of data is available, which is good enough for model identification.

From Figure 3.2, low iron level and high WBC can be observed at the beginning of treatment, whereas EPO dosages are ramping up, meaning EPO is not effective if WBC and iron are not optimal in the body. Similarly, from Figure 3.3, low iron and high EPO level can be observed between 600 to 700 days, where hemoglobin level is decreasing.

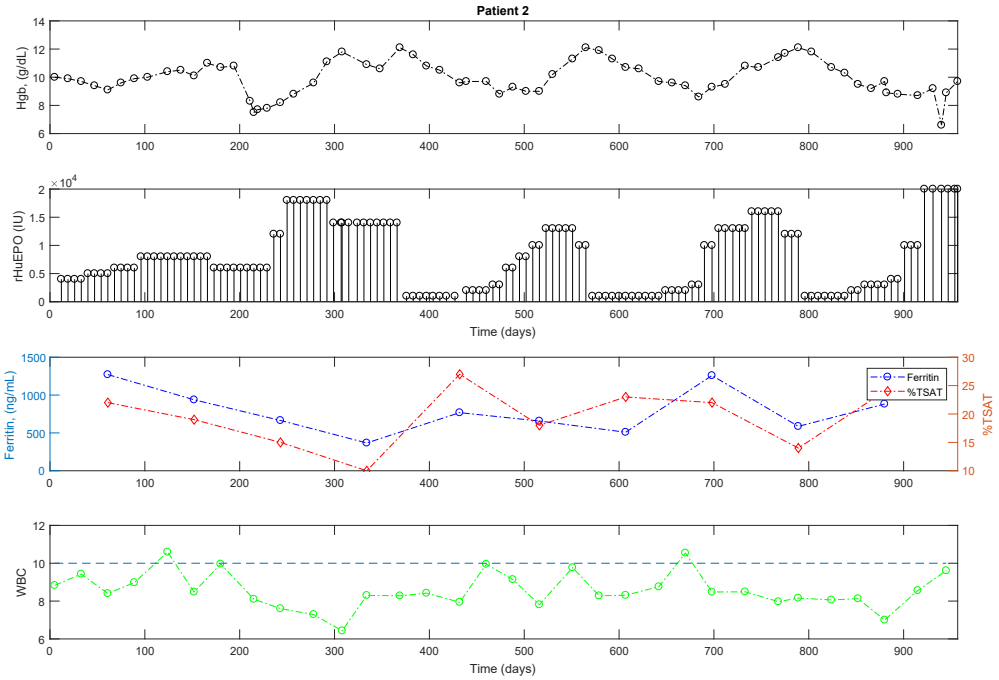


Figure 3.2: Clinical data for patient No.2

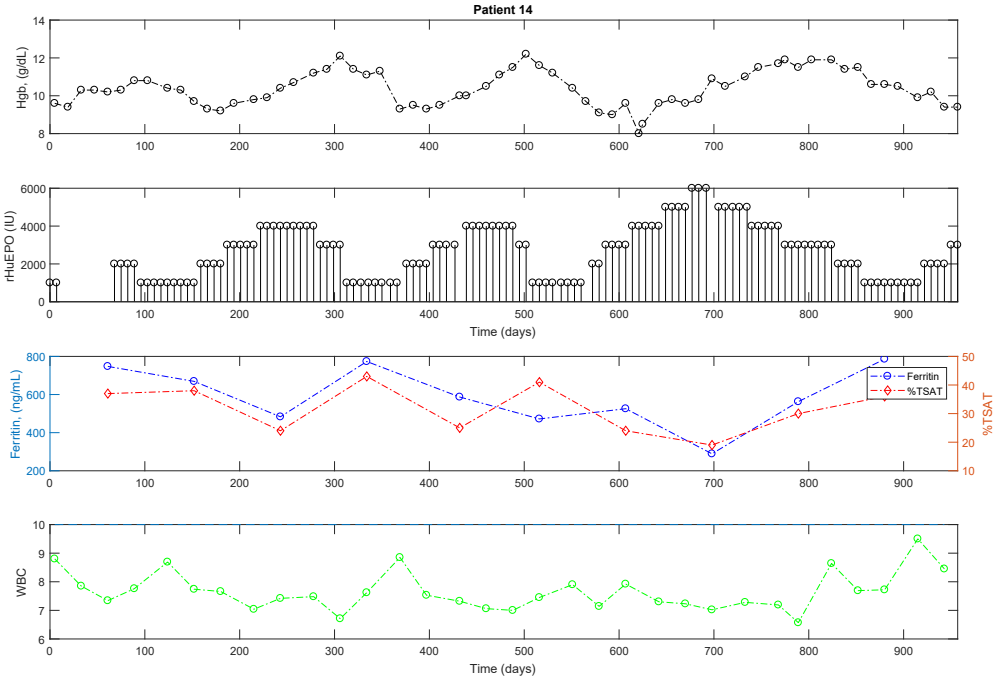


Figure 3.3: Clinical data for patient No.14

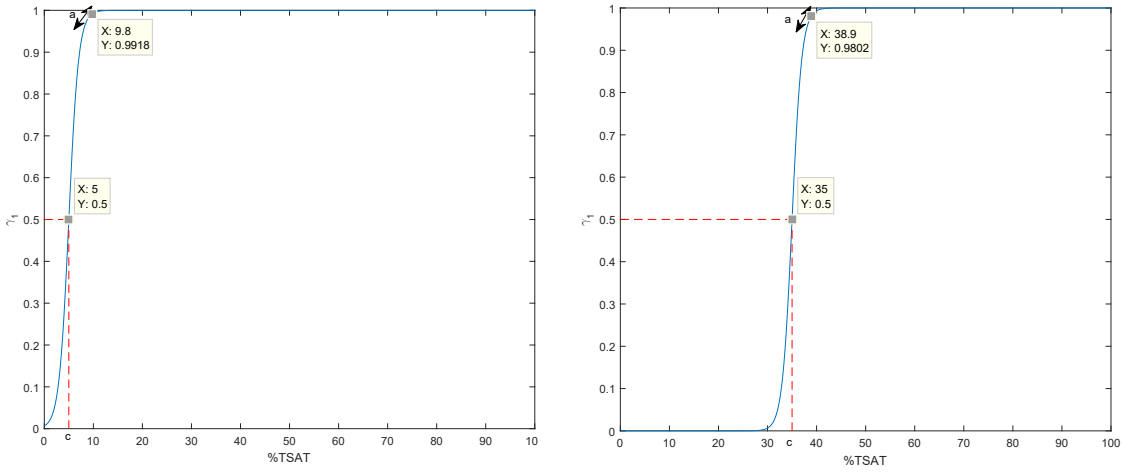
3.2.1 Introducing TSAT in the model

As mentioned in section 3.1.1 erythropoiesis without sufficient iron level in the body do not raise hemoglobin level (Kalantar-Zadeh *et al.*, 2006). According to our methodology, EPO dosage will not be effective in such cases, if TSAT values are lower, and EPO will be moderately effective at moderate TSAT values, and 100% effective if TSAT values are saturated and above the threshold required for the particular patient. This could be understood mathematically with sigmoid function presented in Equation (3.2).

$$\gamma_1 = \frac{1}{1 + e^{-a(\mathrm{TSAT}-c)}} \tag{3.2}$$

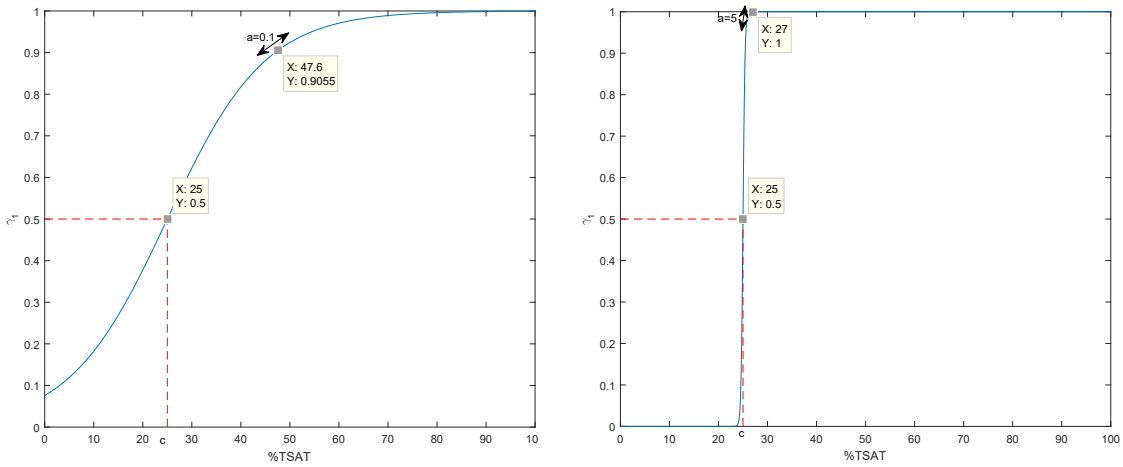
where a = slope of sigmoid function ($0.01 \leq a \leq 5$), c = threshold cut-off when $\gamma_1 = 0.5$ ($5 \leq c \leq 50$). Figures 3.4a to 3.5b shows different sigmoid functions by varying slope and threshold cut-off. Once the γ_1 is calculated, effective EPO is calculated as follow:

$$EPO_{effective} = EPO \times \gamma_1 \tag{3.3}$$



(a) a = 1, c = 5 **(b)** a = 1, c = 35

Figure 3.4: Sigmoid function at different slopes for calculating γ_1 from TSAT



(a) a = 0.1, c = 25 **(b)** a = 5, c = 25

Figure 3.5: Sigmoid function at different slopes for calculating γ_1 from TSAT

Modified C-ARX problem with TSAT

The C-ARX problem developed by J. Ren *et al.*, 2017 with [1 8 1] model order in ARX was used with additional parameter TSAT. The constrained problem is similar to section 2.2, where the ARX model had one 'a' parameter, eight 'b_k' parameters, and one sample delay. The 'b_k' parameters were constrained to follow particular shape. The modified model structure is presented in Equation 3.4 with γ_1 included in optimization problem Equation (3.5).

$$\begin{aligned} Hgb_{t+1} - Hgb_{ss} =& a_1 \ (Hgb_t - Hgb_{ss}) \dots \\ &+ \sum_{k=1}^{8} b_k \ (\gamma_1 EPO_{t-k+1} - EPO_{ss}) + e_t \end{aligned} \tag{3.4}$$

where Hgb_{ss} and EPO_{ss} are steady state hemoglobin and erythropoietin respectively. The overall problem is mixed integer non linear programming problem (MINLP). The objective function is the sum of square errors between actual hemoglobin and proposed hemoglobin level by the model.

$$\min \sum_{t=1}^{tf} [Hgb_t - Hgb_{t,actual}]^2 \tag{3.5a}$$

$$s.t. - Kz_k + 0.001 \leq k - t_{peak} \leq K(1 - z_k) \qquad \forall k = 1, \dots, K \tag{3.5b}$$

$$-M(1 - z_k) \leq \alpha(k-1) - b_k(t_{peak} - 1) \leq M(1 - z_k) \qquad \forall k = 1, \dots, K \tag{3.5c}$$

$$-Mz_k \leq \alpha \exp^{-\beta(k-t_{peak})} - b_k \leq Mz_k \qquad \forall k = 1, \ldots, K \quad (3.5d)$$

$$\gamma_1 = \frac{1}{1 + e^{-a(\mathrm{TSAT}-c)}} \quad (3.5e)$$

$$7.0 \leq Hgb_t \leq 15.0 \;\; \forall t \quad (3.5f)$$

$$7.0 \leq Hgb_{ss} \leq 11.0 \quad (3.5g)$$

$$0.7 \leq a_1 \leq 0.99 \quad (3.5h)$$

$$0 \leq EPO_{ss} \quad (3.5i)$$

$$b_1 = 0 \qquad b_k \geq 0 \;\; \forall k = 2, \ldots, K \quad (3.5j)$$

$$b_k \geq 0.1 \;\; k = k_{peak} \qquad 1.1 \leq t_{peak} \leq 3.9 \quad (3.5k)$$

$$\alpha \geq 0.1 \qquad \beta \geq 0.05 \qquad z_k \in \{0, 1\} \quad (3.5l)$$

Similar to McAllister, 2017, to decrease the computational time, the MINLP was changed to the number of nonlinear programmings (NLP) optimizations. This was performed by all three combinations of z_k that are constrained on t_{peak}. The cost function of all three problems was compared and the lowest one was picked as the best model solution to that particular patient. As mentioned in McAllister, 2017, IPOPT gave more robust solution for the regular C-ARX problem, optimization for modified problems was also done using IPOPT solver. For each patient, first 50% of the data was used as training data set and last 50% as validation data set.

Comparing regular C-ARX and modified C-ARX

Figures 3.6 and 3.7 compares the same clinical patient with regular C-ARX and modified C-ARX. It can be seen that the model performance for modified C-ARX model in Figure 3.7 is 30% better in training and 6% improved in validation portion. Detailed comparisions of results for all the patients are provided in section 3.3. From 137 available patients, 74 patients got improvement in performance in training as well as validation by a minimum of 5%. Average improvement rate for patients is 10%, whereas addition of TSAT did not improve performance for the rest 37 patients. For some patients, improvement rate is higher than 30% and can be seen in Figures 3.8 and 3.9. In regular C-ARX sigmoid function parameters were forced to unity.

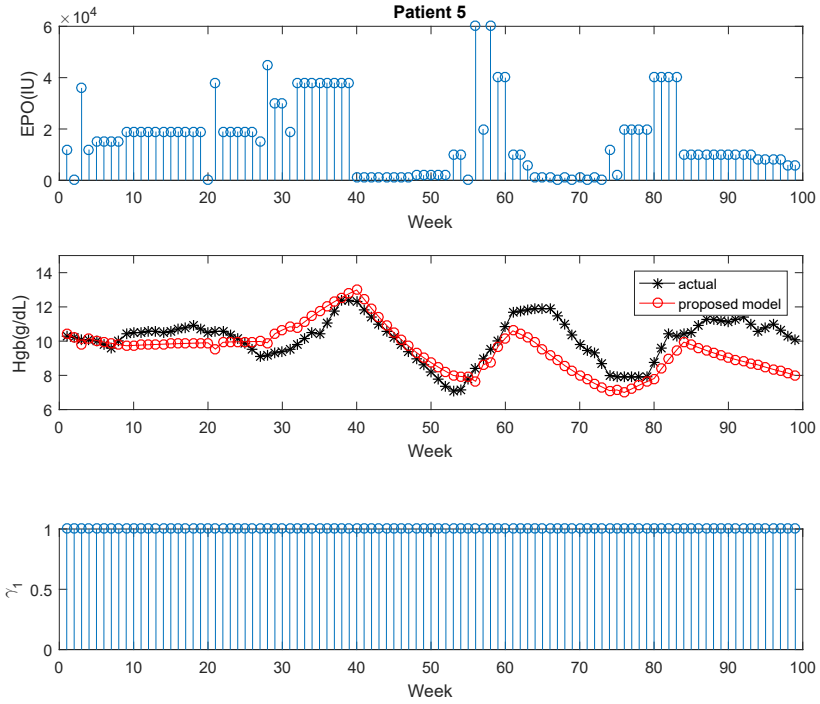


Figure 3.6: Regular C-ARX modeling for patient No. 5

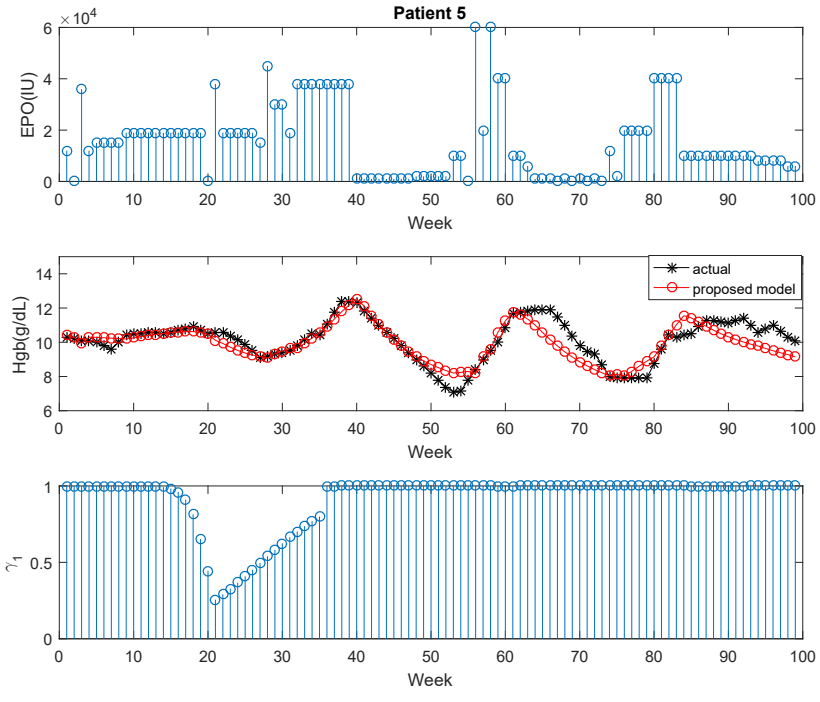


Figure 3.7: Modelfied C-ARX modeling with TSAT for patient No. 5

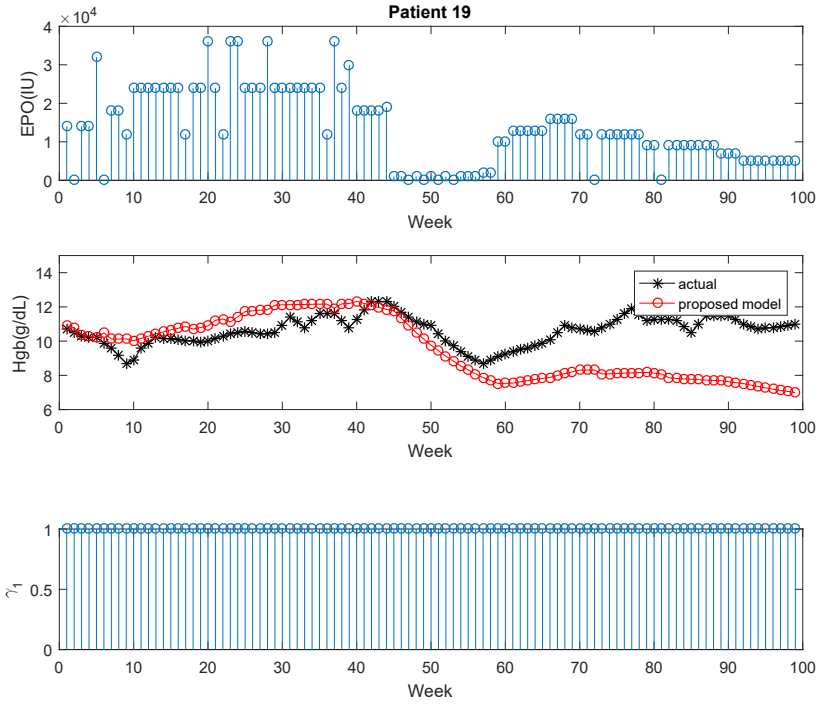


Figure 3.8: Regular C-ARX modeling for patient No. 19

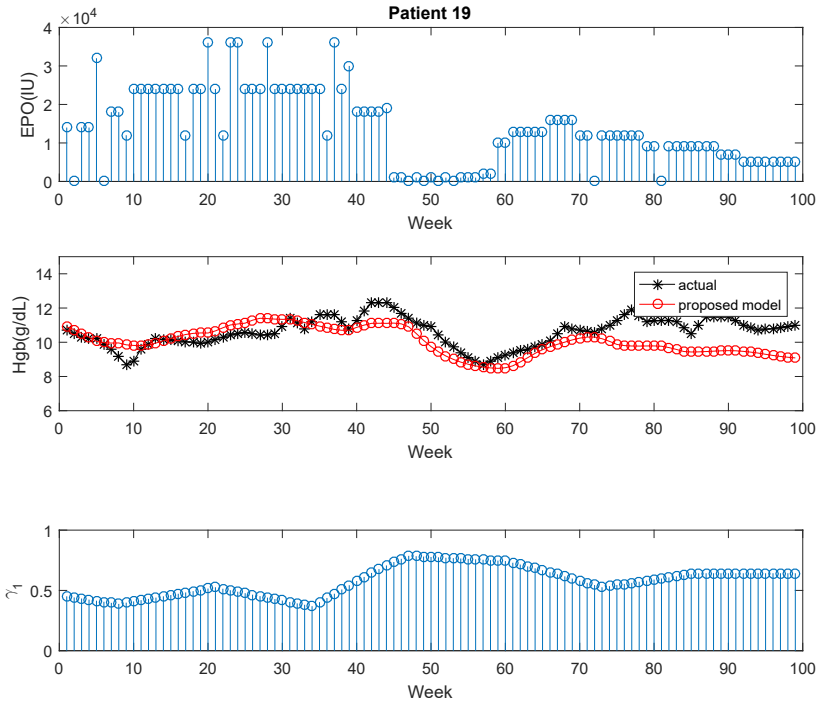


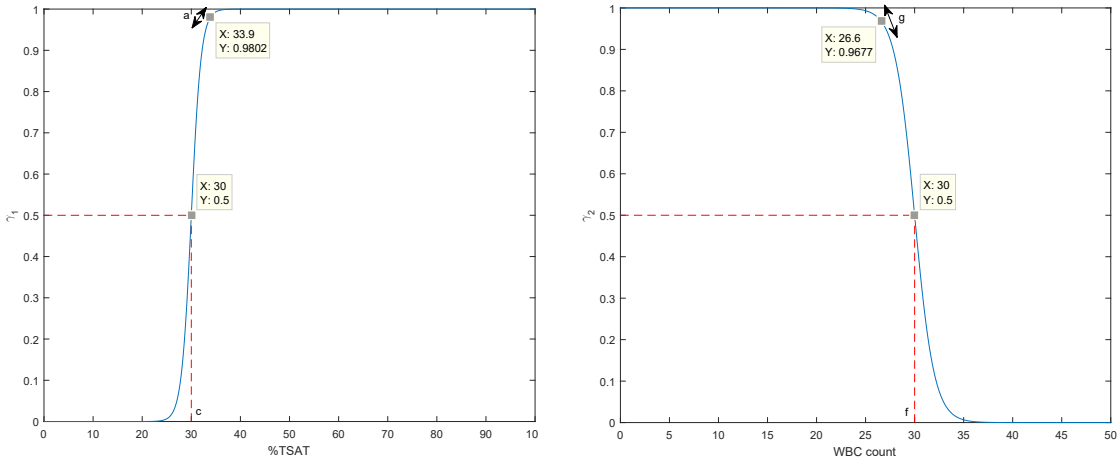
Figure 3.9: Modelfied C-ARX modeling with TSAT for patient No. 19

3.2.2 Introducing TSAT along with WBC

As mentioned in section 3.1.1 and 3.1.3, iron saturation and WBC plays a key role for maintaining hemoglobin in the body. Elevated WBC level represent infection in the body (George-Gay and Parker, 2003). Also, erythropoiesis without sufficient iron level in the body do not raise hemoglobin level (Kalantar-Zadeh *et al.*, 2006). Similar to TSAT, according to our methodology, EPO dosage will not be effective in such cases, if TSAT values are lower, and WBC levels are too high. EPO will be 100% effective if TSAT and WBC are under control for a particular patient. WBC effect could be understand in similar to TSAT effect with negative 'a' parameter and different range for 'c' parameter. This could be understood mathematically with the sigmoid function presented in Equation (3.6).

$$\gamma_2 = \frac{1}{1 + e^{-g(\text{WBC}-f)}} \tag{3.6}$$

where g = slope of sigmoid function $(-5 \leq g \leq -0.01)$, f = threshold cut-off when $\gamma_2 = 0.5$ $(5 \leq f \leq 50)$.



(a) a = 1, c = 30 sigmoid fun for TSAT **(b)** g = -1, f = 30 sigmoid fun for WBC

Figure 3.10: sigmoid function at different cut-offs for calculating γ_1 from TSAT

Figures 3.10a and 3.10b compares sigmoid function for TSAT and WBC side by side. Once the γ_1 and γ_2 are calculated, effective EPO is calculated as follow:

$$EPO_{effective} = EPO \times \gamma_1 \times \gamma_2 \tag{3.7}$$

Modified C-ARX problem with TSAT and WBC

The modified C-ARX problem developed from Equation (3.5), model was further modified to include WBC as an optimization variable. Modeling with TSAT and WBC was performed in two different ways as follow:

1. Simultaneous optimization

2. Sequential optimization

Based on the initial solution provided to the IPOPT solver, either of the optimizations could be solved. Equation (3.8) lists down the calculation of γ_1 and γ_2 respectively.

$$\begin{aligned} \gamma_1 &= \frac{1}{1 + e^{-a(\mathrm{TSAT}-c)}} \\ \gamma_2 &= \frac{1}{1 + e^{-g(\mathrm{WBC}-f)}} \end{aligned} \tag{3.8}$$

Comparing C-ARX, simultaneous, and sequential optimization

For simultaneous optimization, the initial solution from C-ARX was used in the second step to simultaneously optimize the parameters - a, c, g, f; along with the other parameters. The total number of parameters in this case are 18.

For sequential optimization, the initial solution from C-ARX was used in the second step to optimize the parameters a and c simultaneously. The best solution from the second step was then used for the third step to optimize the parameter g and f; along with the other parameters. The total number of parameters in this case are 18. The advantage with this approach is that the solution is always expected to improve in training because of sequential optimization, whereas computational time increases because of the three steps in optimizations.

The modified model structure for C-ARX with simultaneous/sequential optimization could be obtained by including Equations (3.8) in an appropriate fashion. Figures 3.11, 3.12, and 3.13 compare the original C-ARX modeling, simultaneous and sequential optimization problems respectively. It can be seen that the original C-ARX problem has a higher sum of square errors (SSEs) both for the training and validation, whereas simultaneous and sequential optimizations resulted approximately in a 30% improvement in training and 10% and 7% in validation for simultaneous/sequential

optimization respectively. Table 3.2 summarizes the SSE results and computational time for this particular patient.

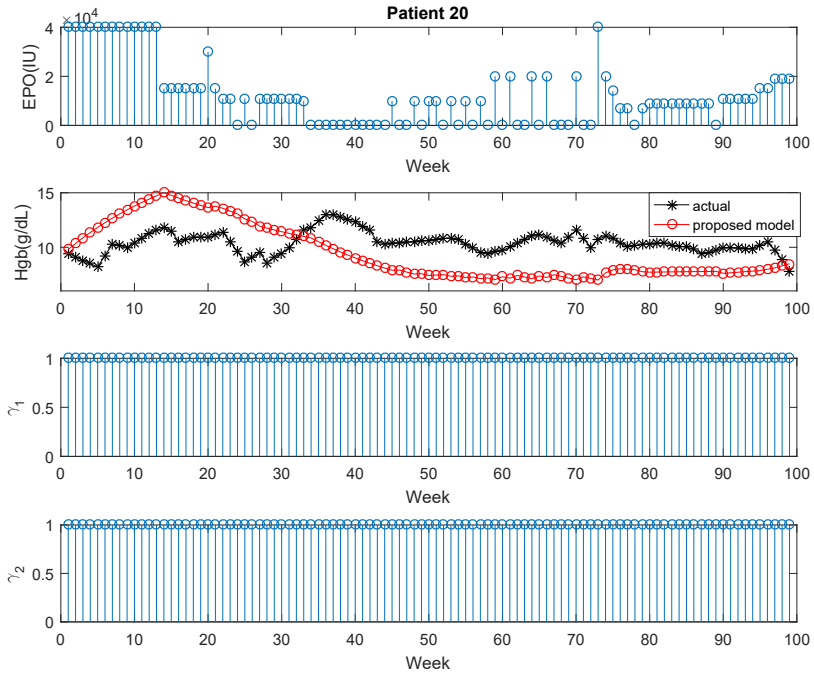


Figure 3.11: Regular C-ARX modeling for patient No. 20

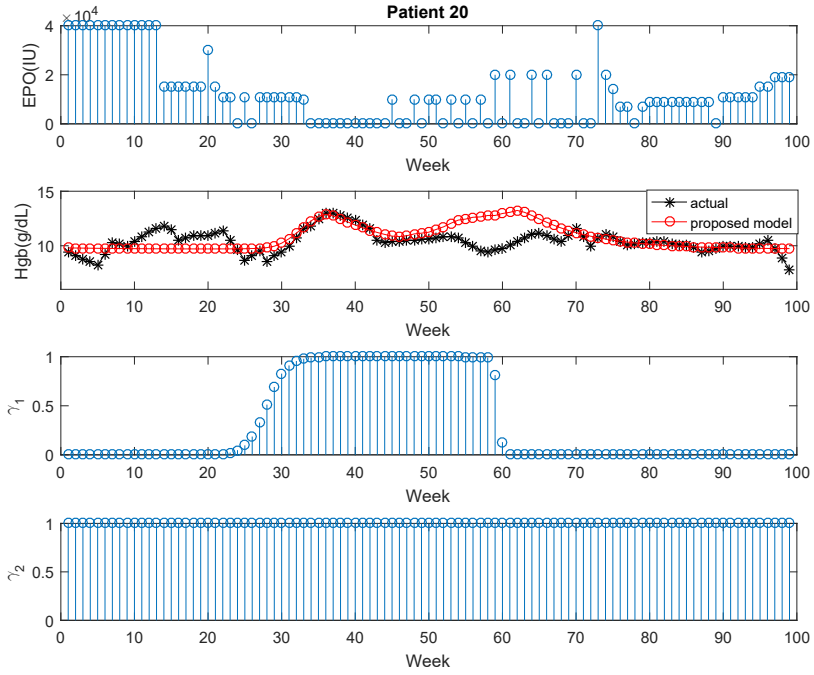


Figure 3.12: Simultaneous C-ARX optimization for patient No. 20

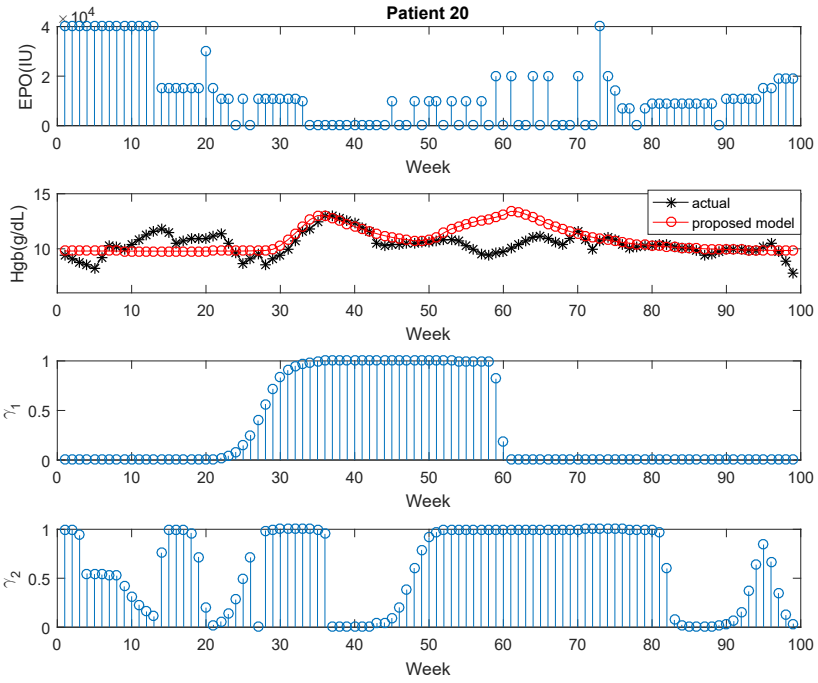


Figure 3.13: Sequential C-ARX optimization for patient No. 20

Table 3.2: Comparision between regular, simultaneous, and sequential C-ARX optimization for patient No. 20

Problem	SSE training	SSE Prediction	Computational time
Regular C-ARX	15.76	9.91	43.28
Simultaneous C-ARX optimization	9.56	8.84	147.99
Sequential C-ARX optimization	9.61	9.14	74.73

It was found that both sequential and simultaneous optimizations with additional parameters (TSAT and WBC) always resulted in a better performance because of the increased degrees of freedom (dof). This improvement varies from patient to patient, where some patients had more than 30% improvement and some patient did not improve and SSE relatively stayed the same as the original C-ARX problem. The increased dof also place models at a risk of overfitting the data. The cumulative results for 137 patients are presented in Table 3.3.

3.2.3 Introducing ferritin along with WBC

As mentioned in section 3.1.1 both TSAT and the serum ferritin are markers for the amount of iron stored in the body. To evaluate which measurement is better for modified C-ARX modeling, optimization with ferritin was also performed. Similar to section 3.2.1 ferritin was used to evaluate the sigmoid function to calculate effective EPO. Modeling was performed for two cases. First, only introducing ferritn in the regular C-ARX problem. Second, simultaneous C-ARX optimization with ferritin and WBC. Results for this section are presented in Table 3.3, where it can be observed that TSAT is a better indicator to be used in the modified C-ARX modeling as compared to ferritin, since performance with TSAT is better for the overall SSE of patients. This could be understood mathematically with the sigmoid function presented in Equation (3.9).

$$\gamma_1 = \frac{1}{1 + e^{-a(\text{ferritin} - c)}} \tag{3.9}$$

where a = slope of the sigmoid function ($0.01 \leq a \leq 5$), c = threshold cut-off when $\gamma_1 = 0.5$ ($5 \leq c \leq 500$). Figure 3.14 show a sigmoid function at a slope of 0.1 and a threshold cut-off of 50. Once the γ_1 is calculated, the effective EPO is calculated as

follows:

$$EPO_{effective} = EPO \times \gamma_1 \tag{3.10}$$

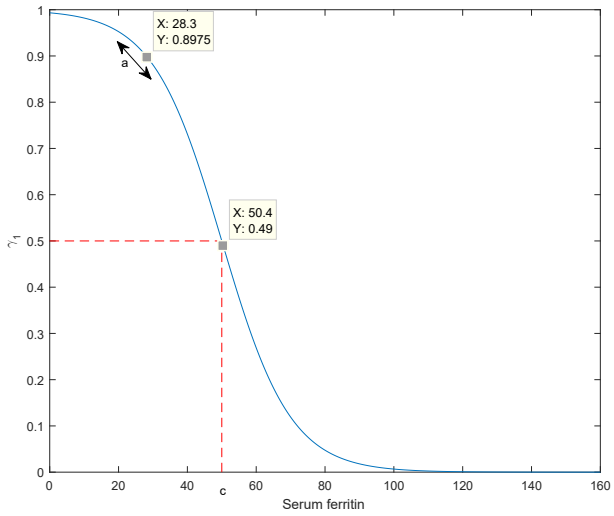


Figure 3.14: Sigmoid function for ferritin with slope 0.1 and threshold 50

3.3 Results

Table 3.3 summarizes the overall modeling results for all the modeling methods that were explored in this chapter.

Table 3.3: Modeling results between different modified versions of the C-ARX models

Modeling method	Training		Validation	
	SSE	% improvement	SSE	% improvement
Regular C-ARX	575.99	–	1002.53	–
Modified C-ARX TSAT	534.68	7.17	1006.98	-0.44
Simultaneous C-ARX TSAT,WBC	512.68	10.99	1017.92	-1.53
Sequential C-ARX TSAT,WBC	513.45	10.85	1050.27	-4.76
Modified C-ARX ferritin	556.40	3.4	1037.34	-3.4
Simultaneous C-ARX ferritin,WBC	560.56	2.68	1025.68	-2.31
Simultaneous C-ARX ferritin,WBC	560.56	2.68	1025.68	-2.31

After performing different modeling techniques, it was found that not a single modeling method gives the best performance for all the patients. Therefore, it was explored how to select the best method for each patient. For this work, training was performed on the first 50% of the data using four modeling techniques: regular

C-ARX, modified C-ARX with TSAT, and modified C-ARX with TSAT and WBC (both simultaneous and sequential). These four methods were then used on the next 25% of the data (validation set) and the best-performed method was used to perform modeling on the last 25% of the prediction set. This will ensure selection of the best modeling technique from all the available methods. On an average computational time increased from 30 seconds (regular C-ARX) to 180 seconds (sequential C-ARX). Tables 3.4 and 3.5 summarize results for the best selected model.

Table 3.4: Modeling results to pick best modified versions of C-ARX models

Modeling method	Training	Validation	No. of models included
Regular C-ARX	610.8	509.53	129
Modified C-ARX TSAT	566.46	532.67	129
Simultaneous C-ARX TSAT,WBC	568.55	566.09	129
Sequential C-ARX TSAT,WBC	559.36	554.57	129
Best picked from validation	–	488.81	129

Table 3.5: Comparision between regular and the best picked modified C-ARX

	Validation		Prediction	
Modeling method	SSE	% improvement	SSE	% improvement
Regular C-ARX	509.53	–	477.5	–
Best picked	488.81	4.07	450.86	5.59

3.4 Conclusions

It was found that TSAT is a more reliable marker for the iron level in the body as it gave better modeling results as compared to ferritin. Model performance increased in training by including parameters such as TSAT and WBC, but validation performance is susceptible to the overfitting of data because of higher dof. It is recommended to perform four different optimizations and pick the best because of the reasonable computational time. It is recommended to use 75% portion of the available data as training and the remaining 25% as the validation data set and choose the best model out of the four, based on validation performance. This methodology ensures the selection of the best model to predict new data.

Chapter 4

Artificial Patient Simulator Design

4.1 Introduction

This chapter focus on patient simulator design. This was identified as a future work in McAllister, 2017 by identifying disturbances which have biological significance. The chapter starts with preliminaries which discuss existing physical hemoglobin model for CKD patients, followed by motivation and physical interpretation of PKPD model parameters, followed by identification of different infections/scenarios that CKD patients can suffer. Later, it was presented how by manipulating the PKPD parameters, tuning process disturbances and measurement noise can relate to change in patient health. The last section of this chapter presents simulated results to compare the developed simulator with the actual patient health. Dataset 1, containing clinical data of 167 patients, was used for this chapter to compare artificial simulator with actual data.

4.1.1 Existing physical model

Physiological relevant mathematical models (PKPD) for hemoglobin response modeling in chronic kidney disease was developed by Y. Chait *et al.*, 2014. Later, J. Ren *et al.*, 2017 simplified and improved the mathematical model using constrained ARX model.

Pharmacokinetic and pharmacodynamics model

Pharmacokinetic is the investigation of the movement of drug inside the body, while pharmacodynamics is the study of the effects that medications have on the body.

The PKPD model developed by Y. Chait *et al.*, 2014 is a continuous time model of delayed differential equations. The model has 8 unique patient-specific parameters. The model parameters are represented in the Table 4.1 (McAllister, 2017, Y.Chait *et al.,* 2014). The system of continuous time delayed differential equation is represented in Equation (4.1).

Table 4.1: Model parameters descriptions for the PK/PD Model

Parameter	Description
H_{en}	Hemoglobin Level due to Endogenous Erythropoietin
μ	Mean RBC life span
V	Maximal clearance rate
K_m	Exogenous erythropoietin level that produces half maximal clearance rate
α	Linear clearance constant
S	Maximal RBC production rate stimulated by E_P
C	Amount of E_P that produces half maximal RBC production rate
D	Time required for EPO-stimulated RBCs to start forming

$$E_{en} = \frac{CH_{en}}{\mu\ K_H\ S - H_{en}} \tag{4.1a}$$

$$\frac{dE(t)}{dt} = \frac{-V\ E(t)}{K_m + E(t)} - \alpha\ E(t)\ +\ dose(t) \tag{4.1b}$$

$$\frac{dR(t)}{dt} = \frac{S\ (E_{en} + E(t-D))}{(C + E_{en} + E(t-D))} - 4\ \frac{x_1(t)}{\mu^2} \tag{4.1c}$$

$$\frac{dx_1(t)}{dt} = x_2(t) \tag{4.1d}$$

$$\frac{dx_2(t)}{dt} = \frac{S\ (E_{en} + E(t-D))}{(C + E_{en} + E(t-D))} - 4\ \frac{x_1(t)}{\mu^2} - 4\ \frac{x_2(t)}{\mu} \tag{4.1e}$$

$$E_p = E(t-D) + E_{en} \tag{4.1f}$$

$$hgb(t) = R(t)K_H \tag{4.1g}$$

where the state $E(t)$ represent the pool of exogenous erythropoietin, $R(t)$ is total population of RBCs within the body, and $x_1(t)$ and $x_2(t)$ are system internal states. E_p is total sum of exogenous and endogenous EPO level. K_H is mean corpuscular hemoglobin, MCH (average amount of hemoglobin per RBC). Range for K_H is 27-33 pg/cell, and an average fixed value of 27.5 pg/cell is used as mentioned in Y. Chait

et al., 2014. Hemoglobin values are calculated from state $R(t)$ by multiplying it with K_H value. Equation (4.2) represents the initial conditions used for the model that have been formulated by McAllister, 2017 that uses previous two Hgb measurements (hgb_1 & hgb_2) at time measurements (t_1 & t_2).

$$\dot{R}_0 = \frac{(hgb_2 - hgb_1)}{K_H(t_2 - t_1)} \tag{4.2a}$$

$$E(0) = 0 \tag{4.2b}$$

$$R(0) = \frac{hgb_1}{K_H} \tag{4.2c}$$

$$x_1(0) = \frac{\mu(H_{en} - \mu\ K_H\ \dot{R}_0)}{4K_H} \tag{4.2d}$$

$$x_2(0) = R(0)\ -\ \frac{4\ x_1(0)}{\mu} \tag{4.2e}$$

4.1.2 Motivation

The artificial patient simulator design is motivated from future work identified in McAllister, 2017, where the simulator design was based on static PKPD model along with data driven noises such as random measurement noise, integrating ramp disturbances, and reoccuring step distrubances. This kind of simulator is hard to explain to a medical professional as it does not relate to biological systems of human health. Also the static PKPD simulator design assumes constant parameters whereas these parameters have physical meaning and do not remain constant. Therefore, static PKPD model can give misleading results. In our work, we also used the physical model (PKPD model) inside simulator. However, instead of keeping the parameters static we considered time-varying parameters that caputres changes in physical condition of patient.

Therefore, to design an efficient simulator which can explain physical disturbances in patient health, the knowledge gap between existing PKPD model and type of diseases that CKD patients suffer needed to be bridged. For this kind of simulator, it might be important to have government endorsement to use of patient simulator as an acceptable tool for pre-clinical trials. This request would be similar to the one for Type 1 Diabetes. The UVA PADOVA Type 1 Diabetes Simulator is an FDA

approved simulator in USA (C. Man *et al.*, 2014), and it replaces animal trials and is sufficient for pre-clinical trials.

4.1.3 Physical interpretation of PKPD parameters

There are eight parameters in the PKPD model. This section describes each of them further with their physical interpretation of the type of effect they can have on patient health.

1. V **- maximal clearance rate (IU/day):**

   Filtration is an important function of kidneys. The analogy for V used in medical terminology is glomerular filtration rate (GFR). Solute filtration/transport decreases as maximal clearance rate decreases (Schwartz and Furth, 2007). Clearance is calculated from excretion rate and plasma concentration by dividing drug elimination with plasma concentration. It reflects the rate at which waste products are cleared from the body by kidney process (Schwartz and Furth, 2007). Total clearance from the body includes renal, salivary, respiratory, etc. Clearance rate measures how fast the body can eliminate drug (Harvison, 2007).

   The mean value of V was found to be 1654.66 IU/day with a standard deviation of 10.24 IU/day. In layman words, higher the V, the faster the drug will be cleared from the body.

2. K_m **- Exogenous EPO level that produces half maximal clearance rate (IU):**

   Also known as ED50 is the drug dose that produces 50% (half-maximal) effect (Toutain, 2002). Stronger the drug, lower the ED50. Different drugs have different kinetics, some of them have first-order kinetics while some have zero-order kinetics.

   For 154 patients, the mean value of Km is 76.52 IU with the standard deviation of 7.32 IU.

3. α **- Linear clearance constant:**

The physical meaning of Alpha is related to V. For 154 patients, the mean value of Alpha is 0.26 with standard deviation of 0.17.

4. S **- maximal RBC production rate stimulated by E_p (cell/day/dL):**

   The physical meaning of S is maximum RBCs that can be produced by the combined effect of exogenous and endogenous EPO level ($E_p = E_n + E$).

   Mathematically, both V and S have the same meaning, but appear in different equations. For 154 patients, the mean value of S is 0.0034 with a standard deviation of 0.0294.

5. C **- Amount of E_p that produces half maximal RBC production rate (IU):**

   Mathematically, C has same the meaning as of K_m but C depends on E_p (= E_n + E), whereas K_m depends only on Exogenous dose.

   For 154 patients, the mean value of C is 22.56 with a standard deviation of 9.78.

6. μ **- Mean RBC life span (Days):**

   As opposed to the adult RBC lifespan of 120 days, the life span of RBC cells produced by erythrocyte is only 60 to 90 (Pearson, 1967). μ represents life-expectancy of RBC cells.

   For 154 patients, mean RBC life span is: 92.84 days with a standard deviation of 23.95 days.

7. D **- Time required for EPO-stimulated RBCs to start forming:**

   D is the total time needed for cells stimulated by progenitor to advance through various stages to fully develop into RBCs (Y. Chait *et al.*, 2014).

   For 154 patients, the mean value of D is 6.33 days with the standard deviation of 0.1059 days.

8. H_{en} **- Hemoglobin Level due to Endogenous Erythropoietin(g/dL):**

   H_{en} is the Hemoglobin Level due to endogenous erythropoietin. In literature, the general range of H_{en} is 4.1 - 9.5 g/dL (Y. Chait *et al.*, 2014).

For 154 patients, the mean value of H_{en} was 6.99 g/dL (slightly on the higher side as compared to literature value) with a standard deviation of 3.01 g/dL. Table 4.2 summarizes PKPD model parameters used for simulator design that are calculated using non-linear least square regression in McAllister, 2017. After outliers removal from 167 patients (where these parameters had negative values i.e, opposite to their physical meaning), 154 patients were considered for further analysis.

Table 4.2: Model parameters range for 154 patients on the PK/PD Model

Parameter	Mean value	Min value	Max value	Standard deviation
V (IU/day)	1654.02	1615.78	1680.45	7.32
K_m (IU)	76.51	35.87	105.92	6.52
α	0.26	0.003	1.28	0.134
S (cells/day/dL)	0.006	0.0003	0.019	0.002
C (IU)	23.03	0.49	92.26	8.31
μ (Days)	92.33	47.68	197.40	18.29
D (Days)	6.33	6.29	6.69	0.050
Hen (g/dL)	7.13	0.59	13.08	2.35

4.2 Identification of disturbances

This section considers various real-life scenarios that can happen to patients and analysis of actual clinical figures for disturbance identification. Table 4.3 summarizes different types of scenarios considered. In clinical figures, Hgb measurement frequency is every two weeks.

Table 4.3: Types of disturbances

Type	Pattern in data	Physical scenario
I	Sharp drop in Hgb / Step change	Blood Loss
II	Hgb remains constant, EPO level increases	EPO resistant
III	Hgb remains constant, EPO level decreases	EPO supporter
V	Hgb ramping down, EPO ramping up	Infections

4.2.1 Blood loss

Blood loss is the type of disturbance when CKD patients loose blood during hemodialysis, internal bleeding, frequent blood testing, etc. In those cases, patient

hemoglobin falls sharply. Analyzing clinical figures step change in hemoglobin was found in patients and following Figures 4.1 and 4.2 are two of such examples. In Figure 4.1 patient suffers a sharp drop in Hgb from 11 to 7 g/dL in 1 week and jumps back in next week. This type of fast recovery in Hgb is rare with only erythropoiesis treatment because the system has delay and it takes D days for cells to progress into RBCs. It was assumed that this type of recovery is possible by other treatment methods such as transfusion.

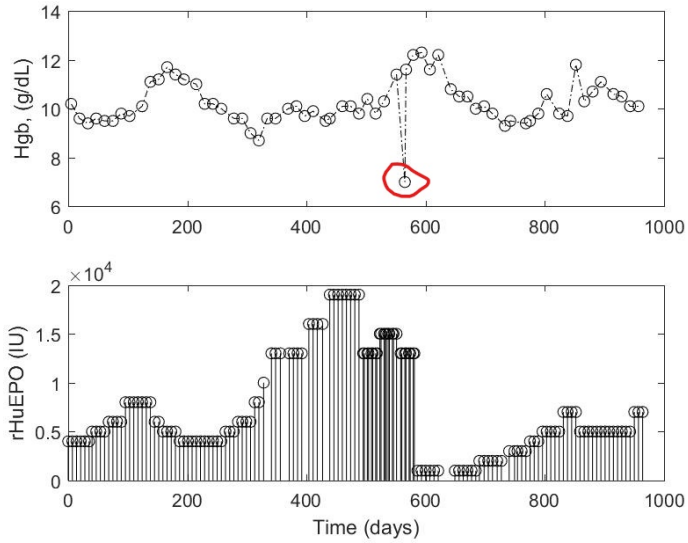


Figure 4.1: Clinical figure of Patient No. 28; blood loss near 600 days

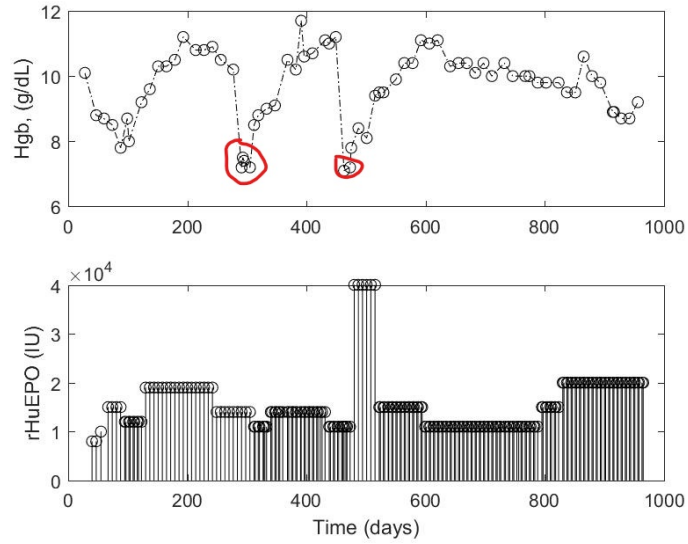


Figure 4.2: Clinical figure of Patient No. 31; two subsequent blood loss scenarios

In Figure 4.1, it can also be observed that patient used less EPO drug dose after 600 days, which reinforced the assumption of patient receiving transfusions. In Figure 4.2, patient suffers two subsequent drops in Hgb at an interval of approximately 200 days. Also highlighted portion of the figure has 3 measurements at the same time, which confirms the sharp drop in Hgb within 2 weeks. After analyzing 167 clinical figures it was found that small blood loses (10 to 25%) drop was quite often, whereas large blood loss up to 25 to 50% drop in Hgb was observed in 3 cases.

It was expected that if PKPD model (Y. Chait *et al.*, 2014) is perfect, it should capture model dynamics as well these step disturbances. Figure 4.3 and 4.4 shows PKPD proposed model against original data. It can be seen that in Figure 4.3 proposed model did not capture step disturbance and in Figure 4.4 model performance near blood loss instance is poor.

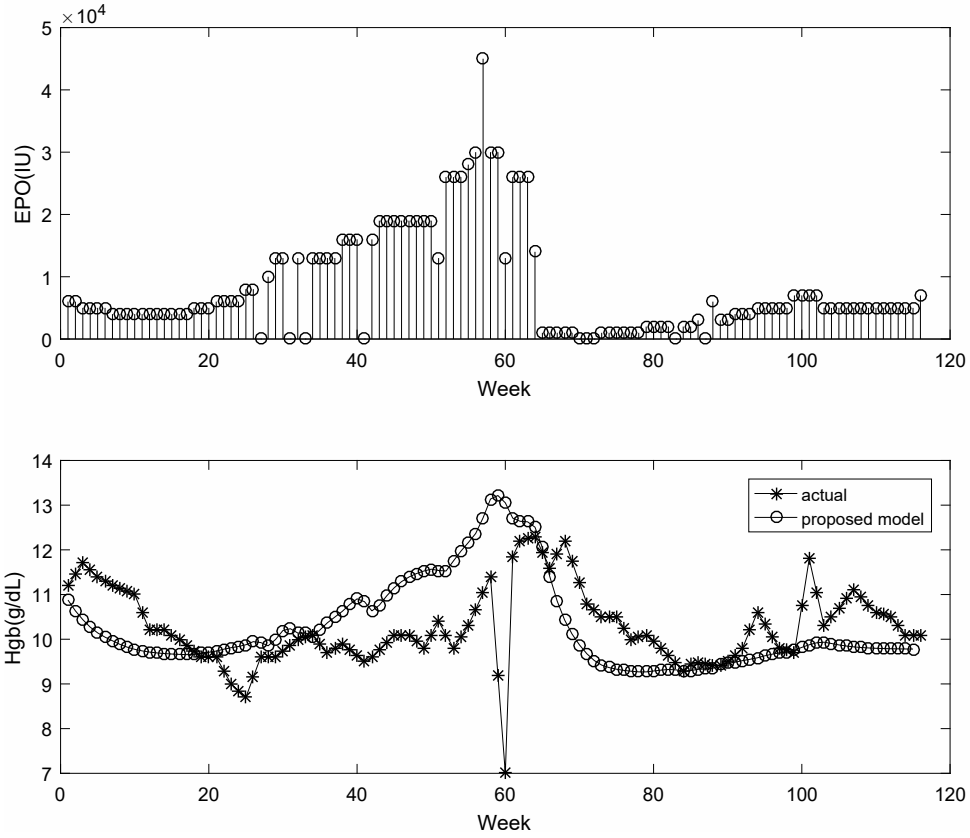


Figure 4.3: Patient 28: PKPD model vs actual data; unable to capture blood losses

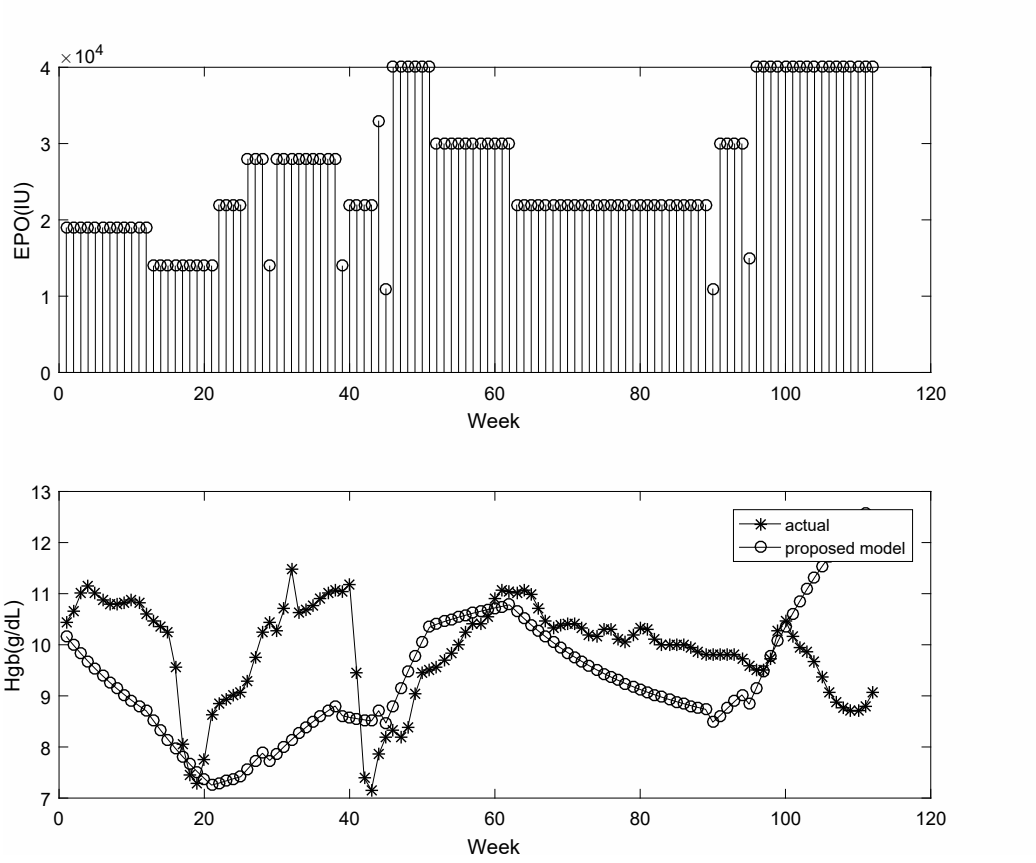


Figure 4.4: Patient 31: PKPD model vs actual data; unable to capture blood losses

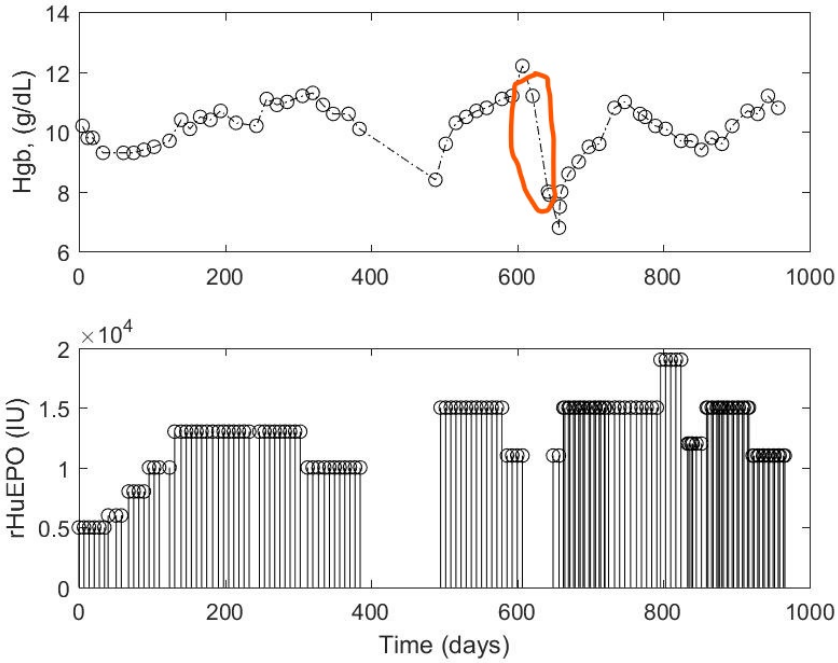


Figure 4.5: Clinical figure of Patient No. 49: not enough measurements

In some cases either enough actual data was not available between Hgb measurements or Hgb drop was slow, those cases were not considered as blood loss, one such example is presented above in Figure 4.5 where at day 616 Hgb is 11.2 g/dL, and then at day 644 Hgb drops down to 8 g/dL.

4.2.2 Drug resistance

It was observed in some clinical figures that the average drug use in first half and later half was quite different. Some patients slowly started using higher average EPO level as compared to before, whereas Hgb levels remained at the same. These types of patients were classified as *drug resistance*.

Figures 4.6 and 4.7, shows two cases of drug resistance, where EPO dose started ramping up after some time and Hgb level stayed relatively the same. Similarly, in Figure 4.7 drug usage in the later half of data is three times higher than initial EPO usage whereas Hgb level oscillates between 9 to 11 g/dL.

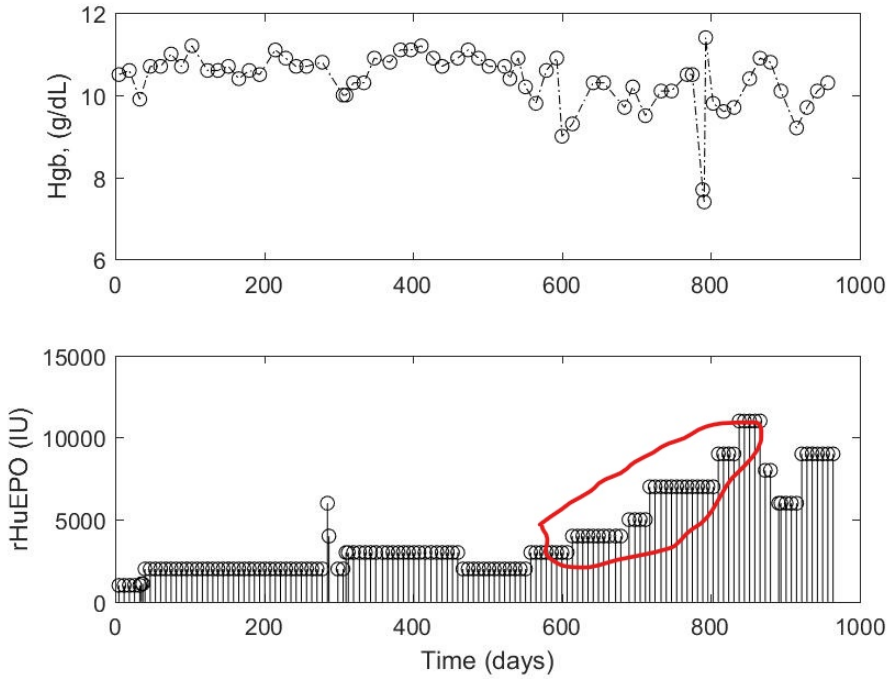


Figure 4.6: Clinical figure of Patient No. 6; drug resistance after 600 days

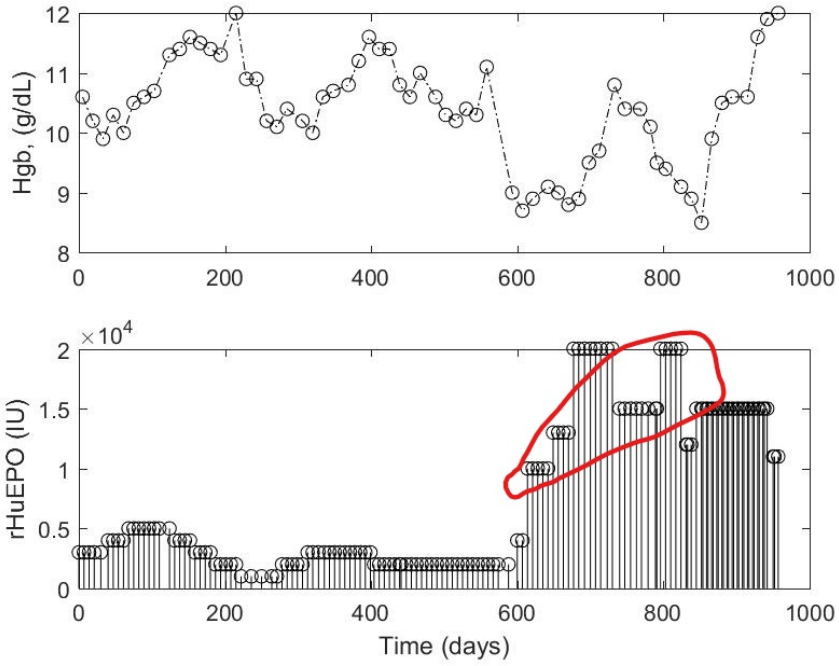


Figure 4.7: Clinical figure of Patient No. 93; drug resistance after 600 days

Figure 4.8 and 4.9 visualize performance of PKPD versus actual data, where regular PKPD have poor performance. It was found that model identified from first 50% data was not able to fit last 50% of validation data

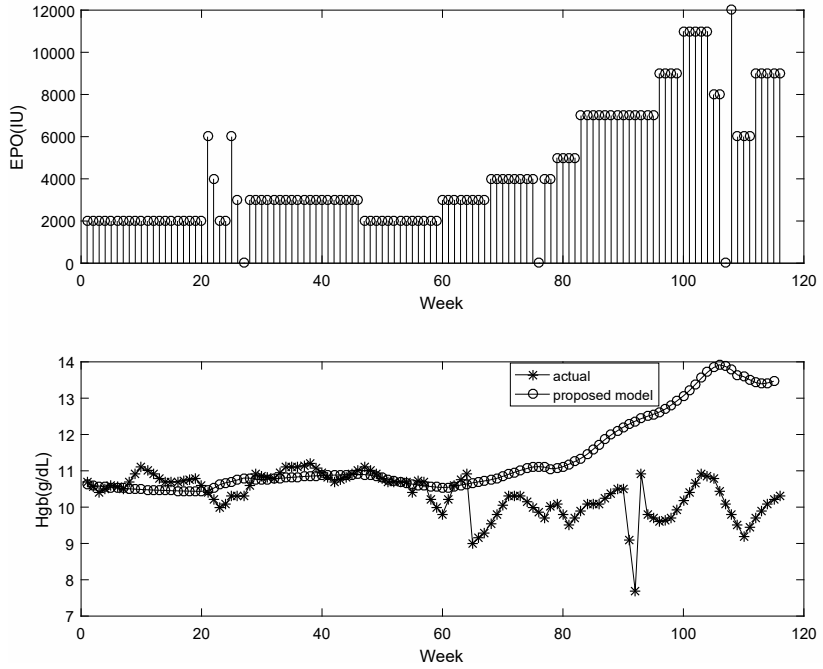


Figure 4.8: Patient 6:PKPD model vs actual data; unable to capture drug resistance

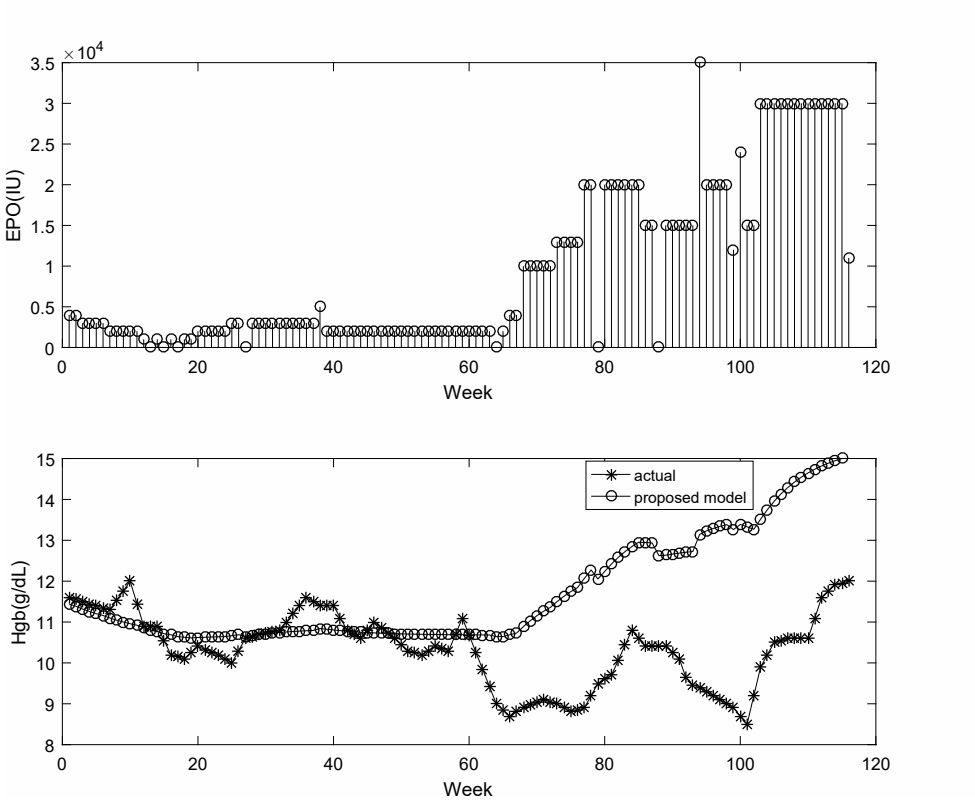


Figure 4.9: Patient 93: PKPD model vs actual data; unable to capture drug resistance

4.2.3 Drug supporter

Opposite to drug resistance scenario, it was observed in some clinical figures that EPO usage significantly decreased after some time. This was attributed to the fact that the patient no longer need a high dosage of the drug. In these cases, EPO starts ramping down but hemoglobin level stayed at the same level. Figures 4.10 and 4.11 explain drug supporter scenarios, where in Figure 4.10 EPO continue ramping down and after 500 days EPO usage is very low. Similarly, in Figure 4.10 patient hgb rise and goes toward higher end and physician started reducing EPO usage.

When regular PKPD optimization was performed on these patients, it was found that regular PKPD trained from initial 50% of data could not capture validation portion. Figures 4.12 and 4.13 summarize modeling figures for EPO supporter scenarios. This ensures that constant parameters values do not capture model dynamics.

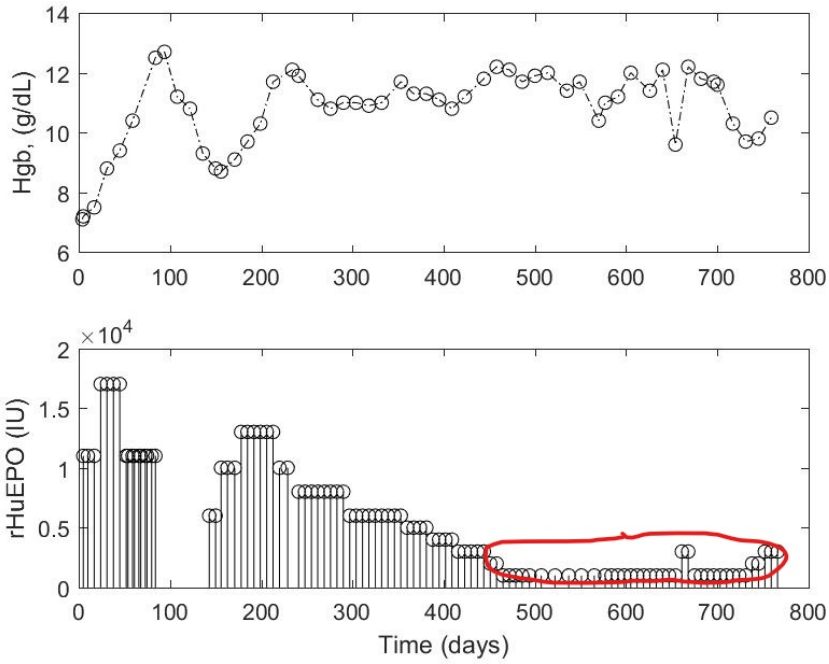


Figure 4.10: Clinical figure of Patient No. 139; drug supporter after 500 days

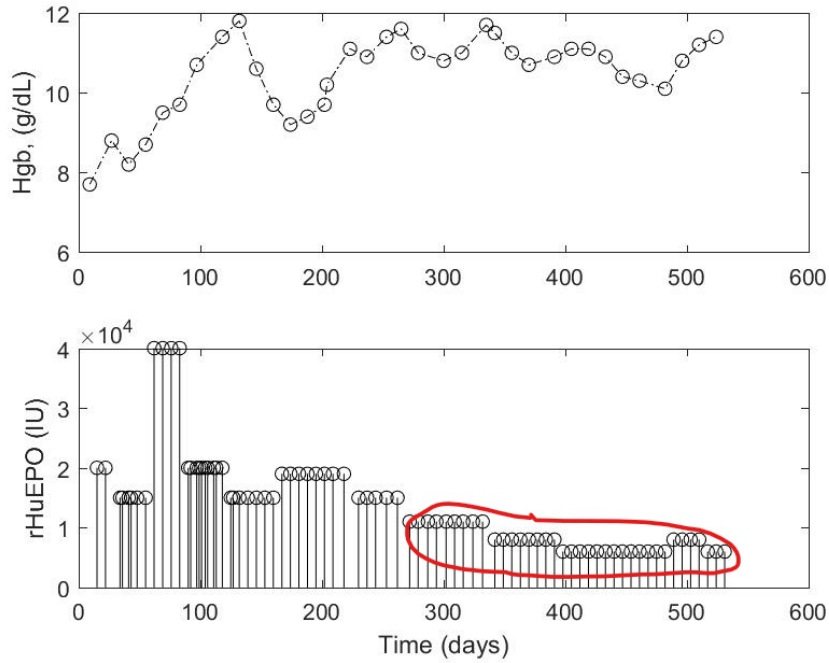


Figure 4.11: Clinical figure of Patient No. 158; drug supporter after 500 days

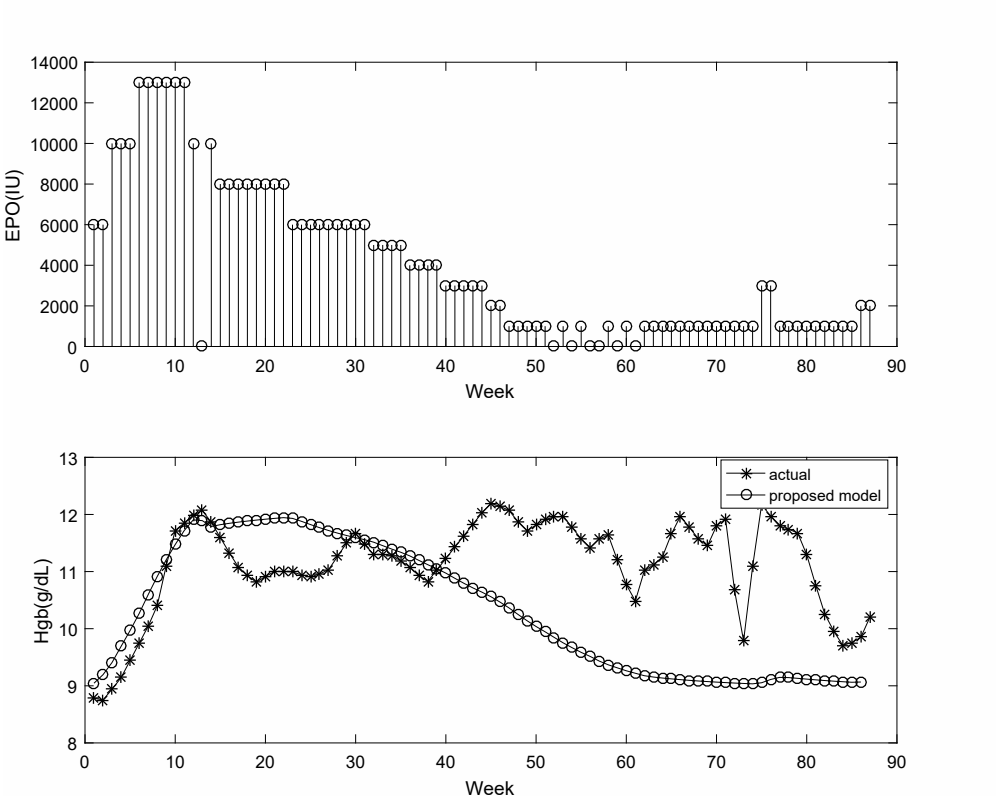


Figure 4.12: Patient 139: PKPD model vs actual data; unable to capture drug supporter

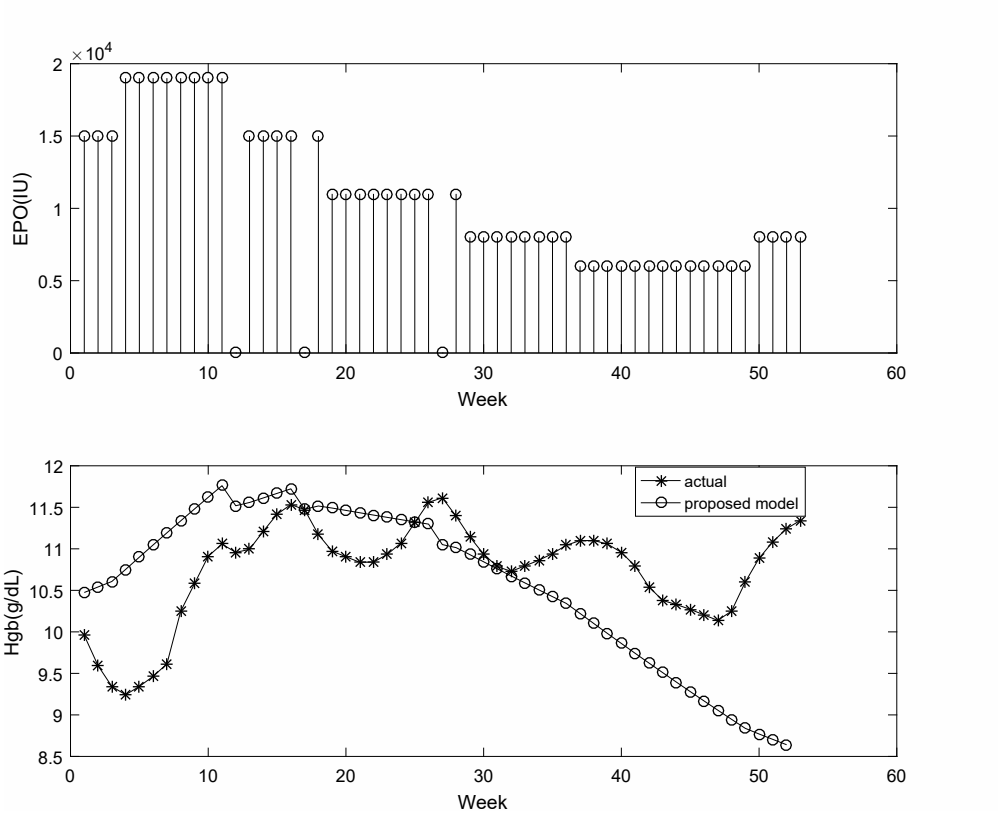


Figure 4.13: Patient 158: PKPD model vs actual data; unable to capture drug supporter

4.2.4 Infections

EPO and Hgb have positive correlation according to PKPD model with the effect of EPO starting with one week delay. But it was observed in some clinical figures that EPO ramps upwards but Hgb ramps downwards, this is opposite to what is expected from drug dosage. After segregating these patients it was also observed that some patients had a longer period where EPO-Hgb had a negative correlation and some patients had a smaller period. These are called *short infections* and *long infections*. Examples of each infection are given below. Short infections are 1 to 3 months long and long infections 3 to 5 months.

Short infections and long infections

Patients who had 1 to 3 months of negative Hgb-EPO correlation, they were classified as short infections. Figure 4.14 is such an example, where the highlighted section had EPO ramping down and Hgb ramping upwards. Figure 4.15 is PKPD modeled data comparison with actual data. Figure 4.16 is a long infection and figure 4.17 compares PKPD with actual data.

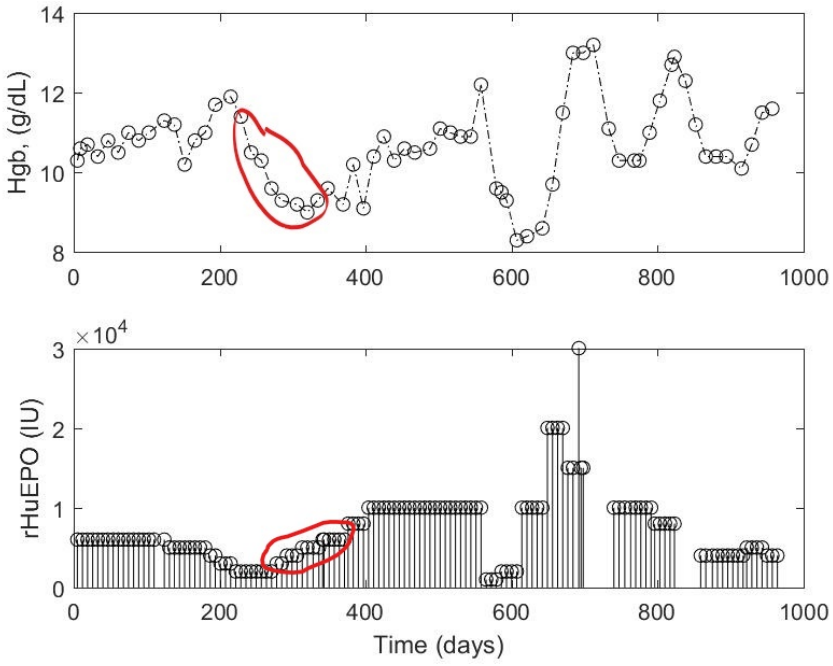


Figure 4.14: Clinical figure of Patient No. 53; short infection between 200 to 400 days

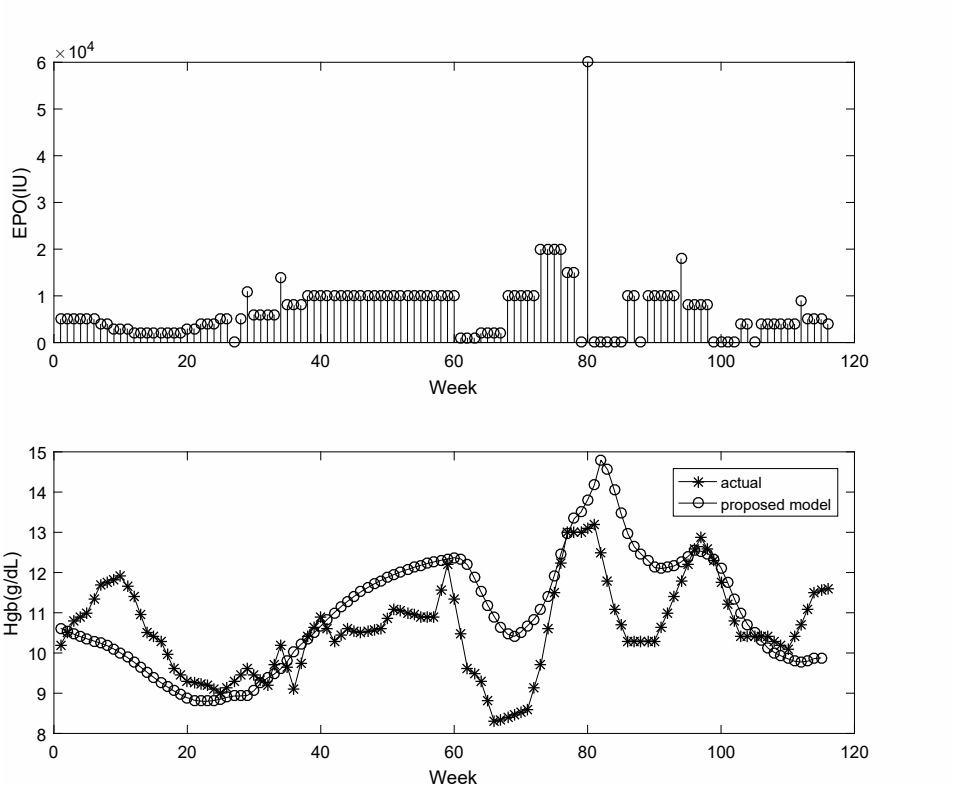


Figure 4.15: Patient No. 53 PKPD model vs actual data unable to capture short infection

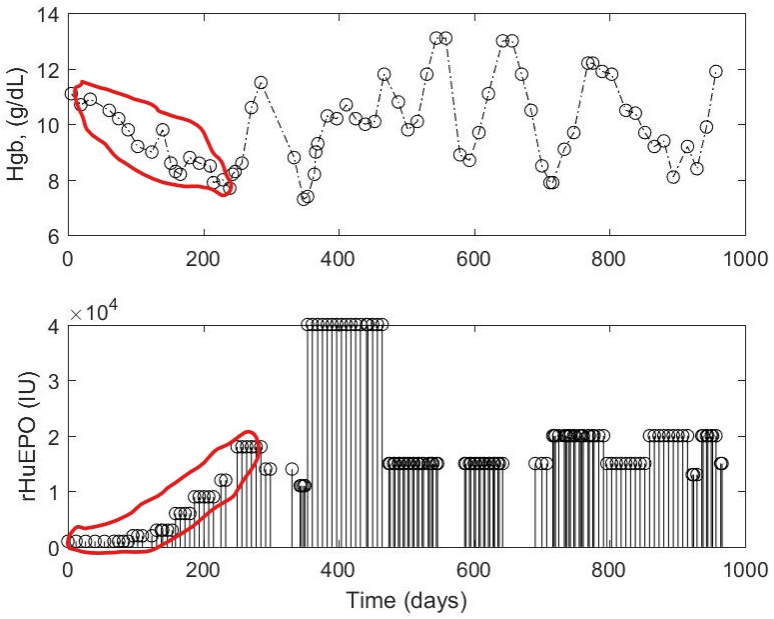


Figure 4.16: Clinical figure of Patient No. 16; long infection between 0 to 200 days

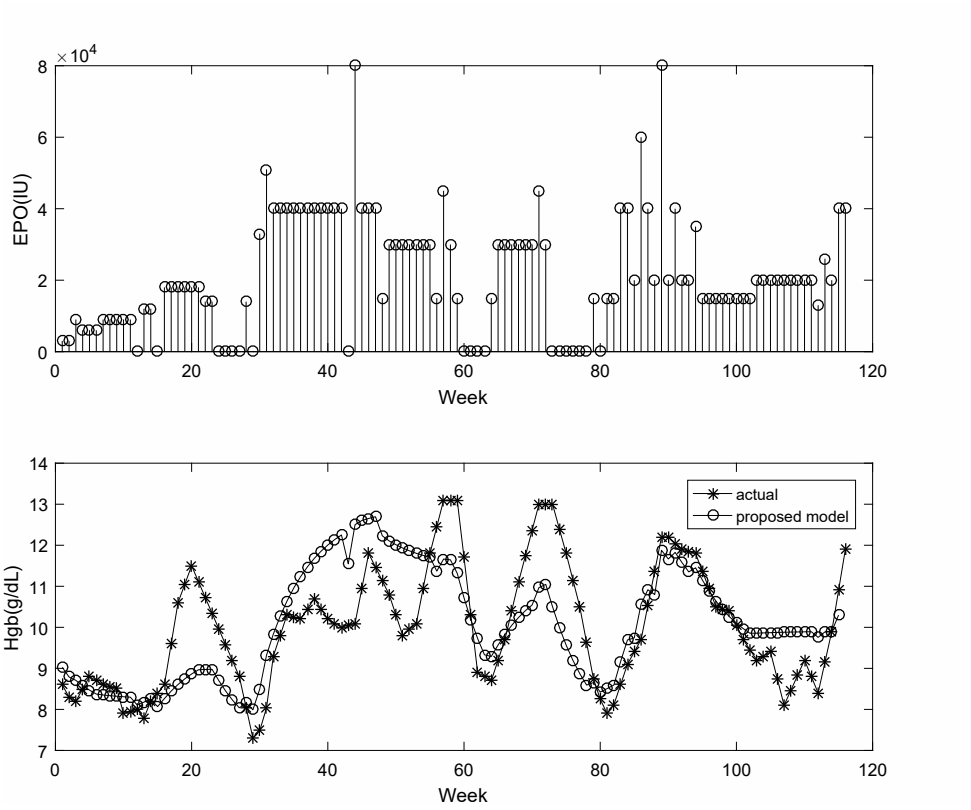


Figure 4.17: Patient 16 PKPD model vs actual data unable to capture long infection

4.3 Simulator design

The patient simulator in following figure 4.18 was designed by incorporating different disturbances and using physician protocol inside the controller.

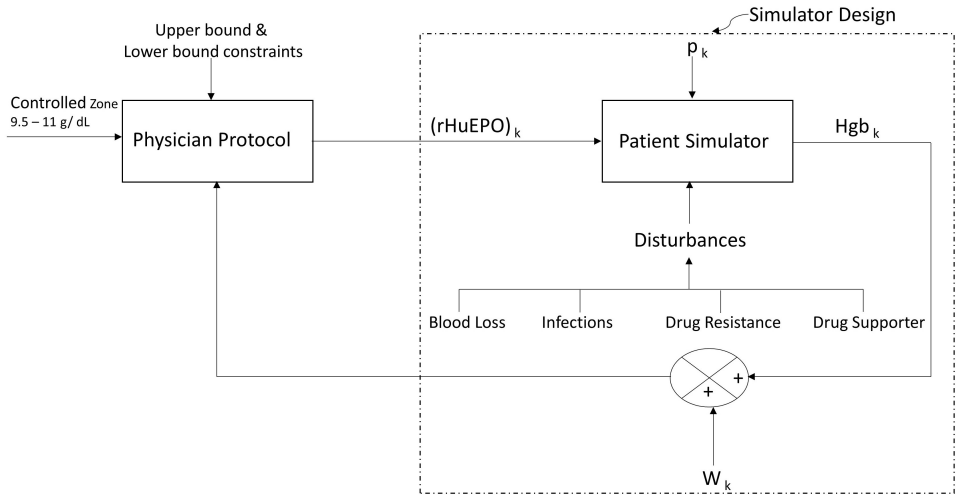


Figure 4.18: Patient simulator design: PKPD model along with disturbances as simulator and physican protocol as controller

PKPD model eight parameters were time-varied to design type of disturbances discussed in previous section. Three type of disturbances were considered while designing the simulator: process disturbance, measurement noise, and change in parameters.

4.3.1 Process disturbance

Although PKPD model is a physical model, it deviates from actual data and has a significant sum of squares square errors. This can be attributed to the fact that the differential equations in Equation (4.1) in PKPD model have some error. The four states in this model are inter-related, therefore, 10% of RBCs from last week was added as p_t to $R(t)$ (Equation (4.3)) which propagates to other states $E(t)$, $x_1(t)$, and $x_2(t)$.

$$\frac{dR(t)}{dt} = \frac{S\ (E_{en} + E(t-D))}{(C + E_{en} + E(t-D))} - 4\ \frac{x_1(t)}{\mu^2} + p(t) \tag{4.3a}$$

$$p(t) = 0.1 \times R(t-1) \tag{4.3b}$$

4.3.2 Measurement noise

Since measurement of Hgb is not perfect and there can be error in meaurement (McAllister, 2017), measurement noise (W_k) was added to Hgb_k which was drawn from $N(0, 0.166^2)$ distribution that correspond to ± 0.5 g/dL.

4.3.3 Change in parameters

Time varying parameters were created that could capture change in patient health as mentioned in section 4.2. This section explains how each disturbance was created. Probabilities of occurance for each type of disturbances were assigned based on actual clinical data and summarized in Table 4.4.

1. **Blood loss**: Blood loss scenarios are created by a step change in RBC population. These disturbances are called as *acute disturbances* (AD), some of them are small (10 to 25% of blood loss) and some of them are a large AD, resulting

into 25 to 50% of blood loss. Analysing the actual clinical data, higher probability was assigned to a small AD than a large AD, as summarized in Table 4.4.

2. **Drug resistance**: These scenarios are created by linearly ramping down S parameter (due to the decrease in maximal RBC production rate stimulated by E_p) and H_{en} parameter (drop in the endogenous level of hemoglobin). Since drug resistance is a slow process, restrictions of occurence was assigned to prevent frequent occurence of drug resistance scenarios.

3. **Drug supporter**: Opposite to drug resistance scenarios, these cases are generated by linearly ramping up S and H_{en} parameters.

4. **Infections**: Short and long infections are attributed to the fact of a decrease in S parameter, or drop in RBC life span (μ parameter), and/or drop in the endogenous level of hemoglobin (H_{en} parameter). All parameters are changed linearly with time (Equation (4.4)) and probabilities of occurence were assigned based on actual clinical data.

$$S = S + S \times N\frac{t - t_2}{t_1 - t_2} \tag{4.4a}$$

$$\mu = \mu + \mu \times N\frac{t - t_2}{t_1 - t_2} \tag{4.4b}$$

$$H_{en} = H_{en} + H_{en} \times N\frac{t - t_2}{t_1 - t_2} \tag{4.4c}$$

where N is change in fractional change in parameter, t_1 and t_2 are start and end time (week) of disturbance, and t is current time (week). In the current simulator after specifying amount of drop required in Hgb and type of disturbance required, change in patient figures can be produced.

Table 4.4: Frequency of occurnace of each scenario

Physical scenario	Probability
Small AD (10 to 25%)	0.10
Big AD (25 to 50%)	0.05
EPO resistant	0.05
EPO supporter	0.05
Short infections	0.25
Long infections	0.05

4.4 Simulation figures

After designing the simulator, figures similar to actual clinical data were obtained and presented below. These figures are close to actual clinical figures and can be compared with different scenarios presented in section 4.2. Figure 4.19 show a blood loss example where a step change was given to RBC population near 500 day and patient suffered blood loss and is recovering slowly after blood loss. Figures 4.20 and 4.21 show drug resistance and drug suppporter examples where average EPO usage in second half is quite different than first half of the data. Figure 4.22 is an simulated patient example who was made to suffer an infection disturbance near 200 day.

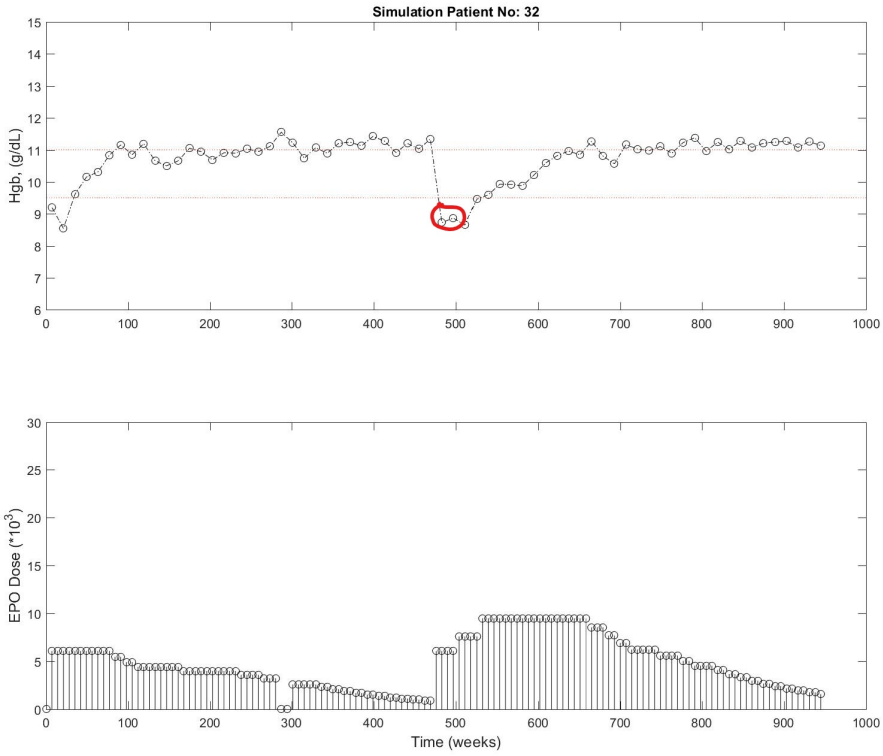


Figure 4.19: Blood loss occuring at day 500 and patient recovering slowly after blood loss

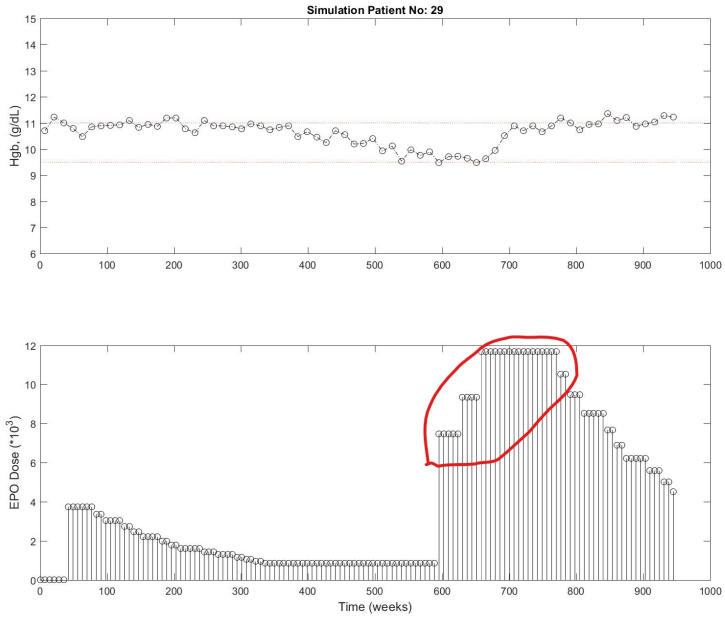


Figure 4.20: Patient suffering from drug resistant after day 600

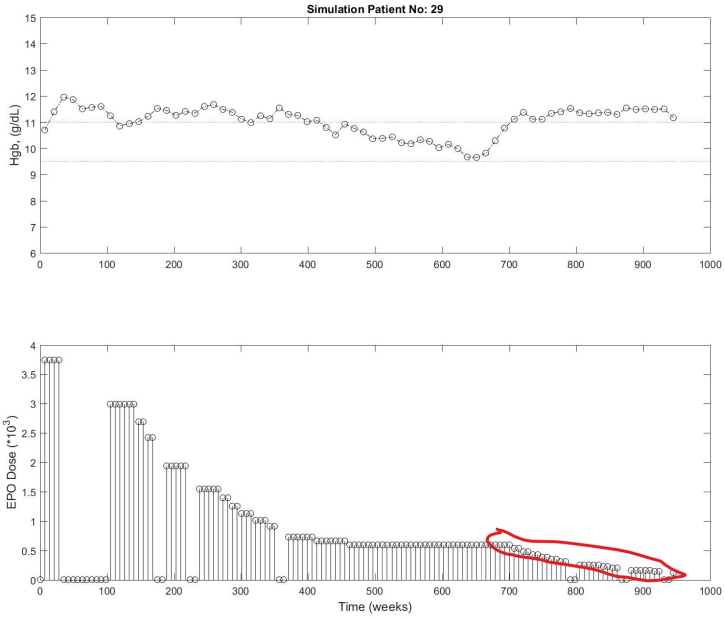


Figure 4.21: Patient recovering and responding as drug supporter in 2nd half of the data

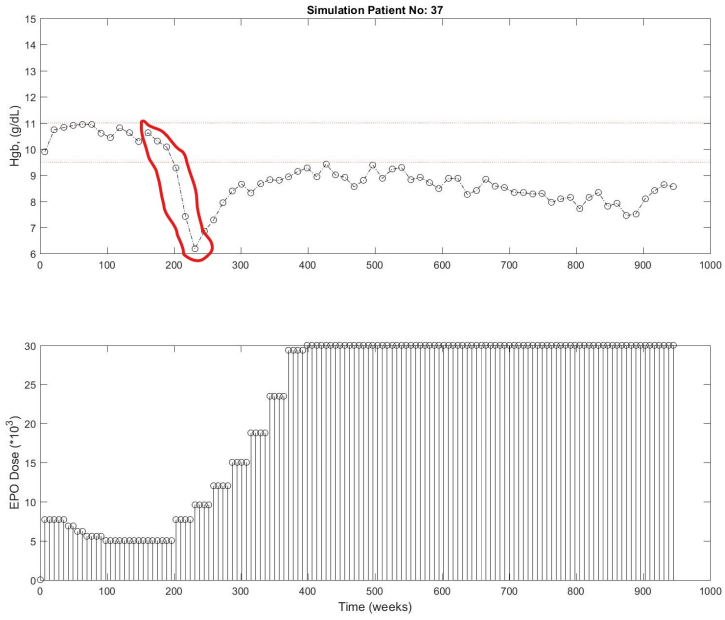


Figure 4.22: Patient suffering with infection near day 200

It was observed in some clinical figures that some patients suffers from more than one above mentioned disturbance, one such example is provided in Figure 4.23 where patient suffers with drug resistance and then recovering back with EPO ramping downwards (drug support scenario).

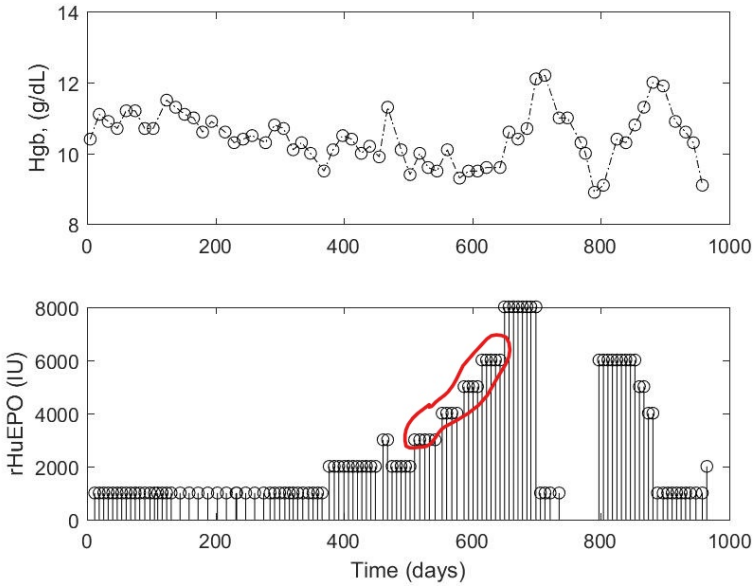


Figure 4.23: Clinical figure: patient suffering drug resistance followed by drug support

Some simulated figures are presented below where mutliple disturbances where added to the patient simulator.

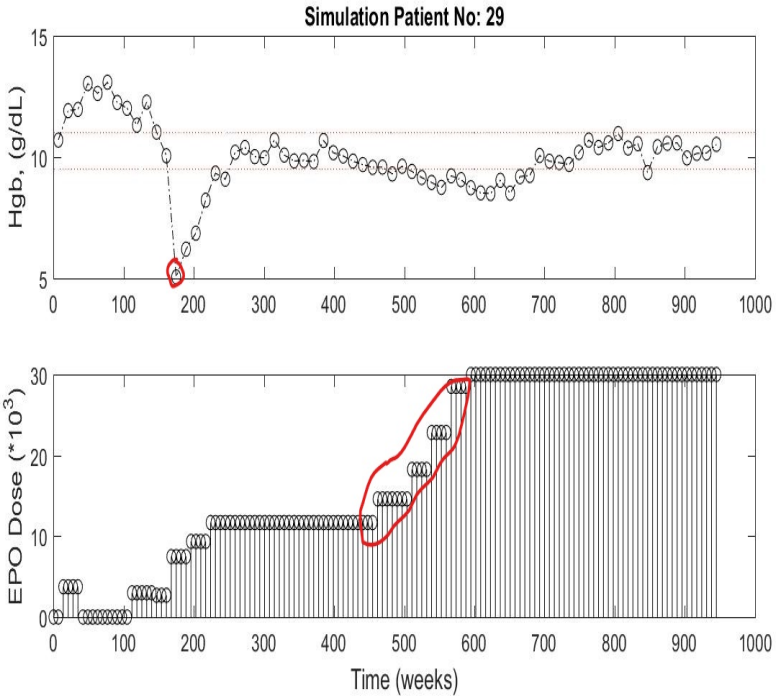


Figure 4.24: Simulated figure: patient suffering blood loss followed by drug resistance

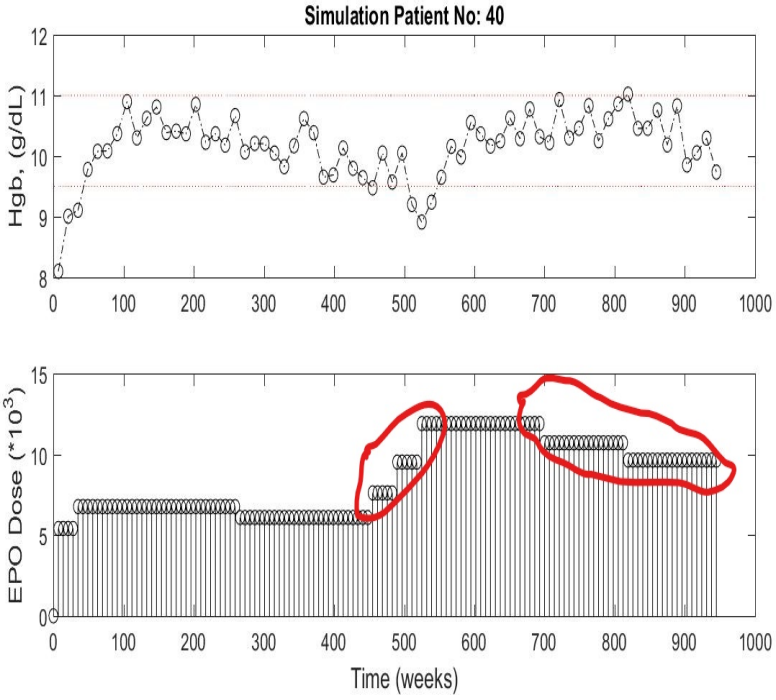


Figure 4.25: Simulated figure: patient suffering drug resistance followed by drug support

4.5 Conclusions

Through this chapter, it was targeted to design an artificial patient simulator which has time-varying parameters to explain the change in patient health. Since parametric disturbances are easy to explain than data-driven disturbances, artificial patient simulator designed with this methodology is more likely to be accepted by medical professionals. After approval from concerned authority, this simulator can be used to rule out ineffective scenarios prior to human use and/or as an equivalent to animal trial.

Chapter 5

Future Work

5.1 Introduction

This chapter summarizes future work identified for this project.

5.1.1 Performance assessment of artificial patient simulator

Testing of artificial patient simulator. The current patient simulator relies on time-varying parameters along with adding process disturbances and measurement noises. Performance assessment of developed artificial patient simulator is the next step and compare performance with data-driven patient simulator developed by McAllister, 2017. Once this comparison is done it can be compared and submitted for FDA approval similar to as it was done for UVA PADOVA Type 1 Diabetes (C. Man *et al.*, 2014).

5.1.2 MIMO model identification

The foundation of CKD hemoglobin model is considering different effects as patients take iron supplements along with EPO. Since patients suffer iron disturbances, it is interesting to see how iron level in the body also varies along with hemoglobin. As explained in section 3.1.2 Gaweda *et al.*, 2014 work motivated our work to explore MIMO system based on iron cycle and erythropoietin cycle in the body, Figure 3.1. Based on preliminary literature review we developed a MIMO model structure. In current data set frequency of measurement of Iron saturation is every other month which we found not appropriate for MIMO model as the iron level in the body could change a lot because of various disturbances suffered by CKD patients. Following

model could be the preliminary multi-input multi-output (MIMO) system structure with TSAT and hemoglobin as system outputs with exogenous EPO and Iron IV as system inputs.

We simplified Gaweda *et al.*, 2014 model and converted into two-input and two-output system, where both systems interact with each other. Figure 5.1 explains the simplified model developed. Model equations are presented in Equations (5.1) and (5.2) where unconstrained ARX model is developed for the MIMO system.

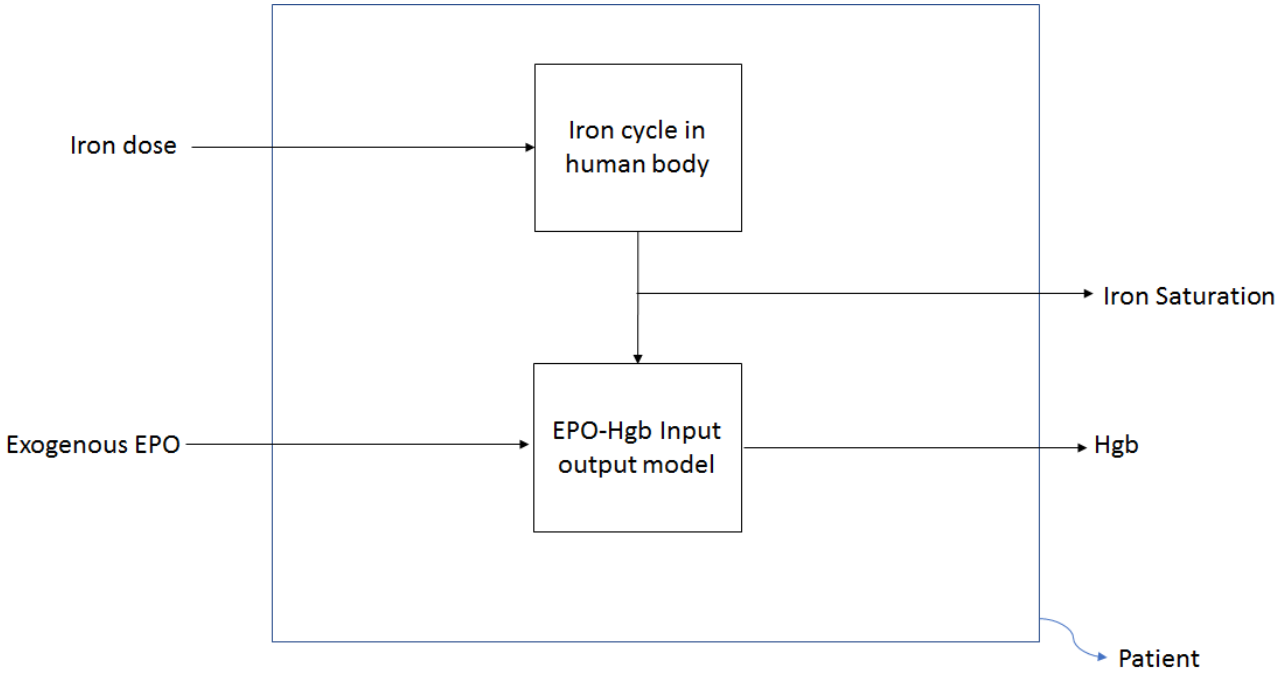


Figure 5.1: MIMO model stucture based on iron cycle and erythropoietin cycle

Output 1: Hemoglobin

$$\begin{aligned} Hgb_{t+1} - Hgb_{ss} =& a_1 \left(Hgb_t - Hgb_{ss}\right) \ldots \\ &+ \sum_{k=1}^{8} b_k \left(EPO_{t-k+1} - EPO_{ss}\right) + c \left(TSAT_t - TSAT_{ss}\right) + e_t \end{aligned} \tag{5.1}$$

Output 2: TSAT level

$$\begin{aligned} TSAT_{t+1} - TSAT_{ss} =& d_1 \left(TSAT_t - TSAT_{ss}\right) \ldots \\ &+ \sum_{k=1}^{8} e_k \left(Iron_{t-k+1} - Iron_{ss}\right) + f_t \end{aligned} \tag{5.2}$$

After receiving appropriate data in which TSAT is measured at a higher frequency, MIMO model structure could be tested.

5.1.3 Using additional parameters

The foundation of any controller is system model. The model that uses more information and analyze different scenarios help decrease uncertainty, increases degrees of freedom, but at the same time possesses the risk of overfitting the data. Similar to TSAT, WBC, ferritin measurements are available for MCH, hematocrit, and urea level in the body. Understanding effect of these parameters could help them including in model identification and improve modeling results. Currently PKPD model uses average value 29.5 for MCH (Y. Chait *et al.*, 2014), a wiser choice will be including interpolated values of MCH.

Bibliography

A. Gaweda, M. Brier, A. Jacobs, G. Aronoff, B. Nathanson and M. Germain (2010). Determining optimum hemoglobin sampling for anemia management from every-treatment data. *Clinical Journal of the American Society of Nephrology* **5**, 1939–1945.

A. Levey, R. Atkins and J. Coresh (2007). Chronic kidney disease as a global public health problem: approaches and initiatives - a position statement from kidney disease improving global outcomes. *Kidney International* **72**, 247–259.

A. Singh, L. Szczech, K. Tang, H. Barnhart, S. Sapp, M. Wolfson and D. Reddan (2006). Correction of anemia with epoetin alfa in chronic kidney disease. *New England Journal of Medicine* **355**, 2085–2098.

Ballabio, Davide and Viviana Consonni (2013). Classification tools in chemistry. part 1: linear models. pls-da. *Analytical Methods* **5**(16), 3790–3798.

C. Man, F. Micheletto, D. Lv, M. Breton, B. Kovatchev and C. Cobelli (2014). The uva/padova type 1 diabetes simulator: New features. *Journal of Diabetes and Technology* **8**, 26–34.

Clements, Michael P and David F Hendry (2005). Evaluating a model by forecast performance. *Oxford Bulletin of Economics and Statistics* **67**(s1), 931–956.

Gaweda, Adam E, Yelena Z Ginzburg, Yossi Chait, Michael J Germain, George R Aronoff and Eliezer Rachmilewitz (2014). Iron dosing in kidney disease: inconsistency of evidence and clinical practice. *Nephrology Dialysis Transplantation* **30**(2), 187–196.

George-Gay, Beverly and Katherine Parker (2003). Understanding the complete blood count with differential. *Journal of PeriAnesthesia Nursing* **18**(2), 96–117.

Group, National Kidney Foundation Anemia Working et al. (2001). Nkf-k/doqi clinical practice guidelines for anemia of chronic kidney disease, 2000. *Am J Kidney Dis* **37**(1), S182–S238.

Harvison, PJ (2007). xpharm: The comprehensive pharmacology reference. *Tenecteplase. Philadelphia*.

J. Ren, J. McAllister, Z. Li, J. Liu and U. Simonsmeier (2017). Modeling of hemoglobin response to erythropoietin therapy through contrained optimization. In: *Proceedings of The 6th International Symposium on Advanced Control of Industrial Processes*. Taipei, Taiwan.

Jauregui, Jeff (2012). Principal component analysis with linear algebra. *Philadelphia: Penn Arts & Sciences*.

Jolliffe, Ian T (1986). Principal component analysis and factor analysis. In: *Principal component analysis*. pp. 115–128. Springer.

Kalantar-Zadeh, Kamyar, Kourosh Kalantar-Zadeh and Grace H Lee (2006). The fascinating but deceptive ferritin: to measure it or not to measure it in chronic kidney disease?. *Clinical Journal of the American Society of Nephrology* **1**(Supplement 1), S9–S18.

Li, Cheng and Bingyu Wang (2017). Fisher linear discriminant analysis.

M. Rosner, W. Bolton (2008). The mortality risk associated with higher hemoglobin: Is therapy to blame?. *Kidney International* **74**, 782–791.

Maitra, Saikat and Jun Yan (2008). Principle component analysis and partial least squares: Two dimension reduction techniques for regression. *Applying Multivariate Statistical Models* **79**, 79–90.

McAllister, Jayson (2017). Modeling and control of hemoglobin for anemia management in chronic kidney disease. Master's thesis. University of Alberta.

Næs, Tormod and Bjørn-Helge Mevik (2001). Understanding the collinearity problem in regression and discriminant analysis. *Journal of Chemometrics* **15**(4), 413–426.

National Kidney Foundation, NKF (2017). Global facts: About kidney disease. *https://www.kidney.org, accessed June 2017*.

Organization, World Health et al. (2011). Serum ferritin concentrations for the assessment of iron status and iron deficiency in populations.

P. Damien, H. Lanham, M. Parthasarathy and N. Shah (2016). Assessing key cost drivers associated with caring for chronic kidney disease patients. *BioMed Central Health Services Research*.

Pearson, Howard A (1967). Life-span of the fetal red blood cell. *The Journal of pediatrics* **70**(2), 166–171.

Schwartz, George J and Susan L Furth (2007). Glomerular filtration rate measurement and estimation in chronic kidney disease. *Pediatric nephrology* **22**(11), 1839–1848.

Singh, A. (2007). The target hemoglobin level in patients on dialysis. *Dialysis and Transplantation* **36**, 1–3.

Suryanarayana, TMV and PB Mistry (2016). Principal component analysis in transfer function. In: *Principal Component Regression for Crop Yield Estimation*. pp. 17–25. Springer.

Toutain, Pierre-Louis (2002). Pharmacokinetic/pharmacodynamic integration in drug development and dosage-regimen optimization for veterinary medicine. *Aaps Pharmsci* **4**(4), 160–188.

V. Jha, G. Garcia-Garcia and K. Iseki (2013). Chronic kidney disease: Global dimension and perspectives. *Lancet* **382**, 260–272.

Varmuza, K. and P. Filzmoser (2016). *Introduction to Multivariate Statistical Analysis in Chemometrics*. CRC Press.

W. Couser, G. Remuzz, S. Mendis and M. Tonelli (2011). The contribution of chronic kidney disease to the global burden of major noncommunicable diseases. *Kidney International* **80**, 1258–1270.

Walters, GO, FM Miller and M Worwood (1973). Serum ferritin concentration and iron stores in normal subjects. *Journal of Clinical Pathology* **26**(10), 770–772.

Wold, Svante, Michael Sjöström and Lennart Eriksson (2001). Pls-regression: a basic tool of chemometrics. *Chemometrics and intelligent laboratory systems* **58**(2), 109–130.

Xiong, Tao and Vladimir Cherkassky (2005). A combined svm and lda approach for classification. In: *Neural Networks, 2005. IJCNN'05. Proceedings. 2005 IEEE International Joint Conference on*. Vol. 3. IEEE. pp. 1455–1459.

Y. Chait, J. Horowitz, B. Nichols, R. Shrestha and M. Germain C. Hollot (2014). Control-relevant erythropoiesis modeling in end-stage renal disease. *IEEE Transactions on Biomedical Engineering* **61**, 658–664.

Z. Jing, Y. Wei-jie, Z. Nan, Z. Yi and W. Ling (2012). Hemoglobin targets for chronic kidney disease patients with anemia: A systematic review and meta-analysis. *Journal of American Institue of Chemical Engineers* **7**, 1–9.