Incorporating Prior Information to Enhance Quantitative ELISA

Sarah Marie Reehl

Follow this and additional works at: https://digitalcommons.usu.edu/gradreports

Recommended Citation
Reehl, Sarah Marie, "Incorporating Prior Information to Enhance Quantitative ELISA" (2015). All Graduate Plan B and other Reports. 495.
https://digitalcommons.usu.edu/gradreports/495
INCORPORATING PRIOR INFORMATION TO ENHANCE QUANTITATIVE ELISA

by

Sarah M. Reehl

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

MASTER OF SCIENCE

in

Industrial Mathematics

Approved:

James Powell
Major Professor

Luis Gordillo
Committee Member

Guifang Fu
Committee Member

Mark MacLellan
Vice President for Research and Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah
2015
Abstract

Incorporating Prior Information
to Enhance Quantitative ELISA

by

Sarah M. Reehl, Master of Science
Utah State University, 2015

Major Professor: Dr. James Powell
Department: Mathematics and Statistics

The enzyme-linked immunosorbent assay (ELISA) is a powerful quantitative tool with predictive abilities completely determined by the accuracy of a five-parameter curve fit to an analyte standard. Parameterizing the standard uses a weighted sum of squares (wSSE) on data defined in a serial dilution using 12-16 out of 96 available wells. The wells used for the standard are wells lost to testing samples, inferring cost on the ELISA consumer and limiting the amount of standard data for fitting to two replicates at each concentration. In addition to limited standard wells, parameterization of the standard is highly subject to fitting error resultant from outlying data, which is impossible to detect with only two replicates, and spurious minima surrounding optimal parameters. If the standard curve cannot properly approximate samples of known concentration, then the inversion of the standard will not accurately quantify samples of unknown concentration. Both inaccurate parameterization and lengthy standard dilution series then add unnecessary expense to ELISA consumers. Our idea is to address outliers, parameterization improvements, and the design of standard dilutions using a Bayesian perspective to introduce prior information.

First, using data from a variety of ELISA runs, we will build prior distributions
by fitting the five-parameter curve to “good” data using wSSE. Distributions of credible responses at particular concentrations will be developed by sampling the priors and evaluating the five parameter curve at those samples. Outliers can then be characterized in an analysis of the tails of the posterior response distribution. An appropriate credible level will be developed for a particular analyte and then tested against others. In parameterization, we will avoid spurious local error minima by maximizing the posterior likelihood for each analyte. A sensitivity analysis will determine a weighting for prior parameter distributions that maintain constant information in the face of new data.

We then remodel the standard design, utilizing fewer standard sample concentrations. We also test alternate standard designs by removing dilution concentrations sequentially from the weighted likelihood and testing a coefficient of determination using all available data. The design which causes the least degradation to the coefficient of determination across all replicate series of all analytes is the proposed design. For removals of three to five dilutions we assess the performance of the standard parameterization by the percentage of acceptably fit data and the percentage of replicate series with acceptably fit data in the low concentration spectrum. We show assessments for parameterization both with and without prior information. Finally, we comment on further development of the priors, problems in weighting and transforming data, and predicting unknown concentrations without using a reference model.

(49 pages)
The enzyme-linked immunosorbent assay (ELISA) is a powerful quantitative tool with predictive abilities completely determined by the accuracy of a five-parameter curve fit to an analyte standard. Parameterizing the standard uses 12-16 out of 96 available wells that are then lost to testing samples, inferring cost on the ELISA consumer and limiting the amount of standard data for fitting to two replicates at each concentration. In addition to limited standard wells, parameterization of the standard is highly subject to fitting error resultant from outlying data, which is impossible to detect with only two replicates, and spurious minima surrounding optimal parameters. If the standard curve cannot properly approximate samples of known concentration, then the inversion of the standard will not accurately quantify samples of unknown concentration. Both inaccurate parameterization and lengthy standard dilution series add unnecessary expense to ELISA consumers. Using data from a variety of ELISA runs, we will address outliers, poor parameterization, and standard design efficiency using a Bayesian perspective to introduce prior information and influence the probability of parameters in fitting the five-parameter model.

The prior information we utilize is the data collected on an analyte during quality control checks of plate development. Parameterizing all of the quality control replicates
allows us to develop a collection of best-fit parameters and summarize a likely parameter
distribution by mean and variance. In effect, the priors cause the curve fitting algorithm
to select parameters that are similar to past parameterization instances rather than
parameters that are influenced only by the current data set and whatever problems it
might have.

First, we use prior information to detect outliers. We sample at random from the
prior distribution and use the sampled parameters to form a collection of hypothetical
standard curves. At any given concentration these simulations give a population of
response pixel intensities where we can exclude intensities that occur outside of a given
frequency boundary. Our results indicate that a boundary near 0.05 tends to remove
few data points while improving measures of fit.

Next, the prior information is weighted to correct for the negative effects of poor
quality data. We add an extra parameter to control the weighting of prior parameter
probability and penalize for the selection of parameters that are unlikely according to
prior information. The impact of the poor quality data on the parameterization of the
curve is lessened by using more prior information. This approach also avoids the spurious
minima that occur in the curve-fitting algorithm.

Lastly, we use prior probabilities to guide parameterization using fewer standard
wells in an effort to save on cost to the consumer. We assume a seven-part serial
dilution and skip those dilutions whose removal degrades the curve fit the least across
all analytes tested. We are able to reduce the number of wells used by the standard,
without serious detriment to the curve fit, from sixteen wells to eight in total. This
standard design requires the use of prior information to guide parameter selection, but
in return provides the consumer with the ability to test more samples per plate. We
formulate three different designs across analytes that use ten, eight, or six wells to
determine the standard, recommending the eight point design.

Using prior information requires relatively little effort. Standard quality control
efforts produce the necessary prior information and the only information needed from
the priors are parameter means, standard deviations, and a priori weighting schemes.
In return, the consumer can easily detect outliers, improve parameterization results, and use fewer wells to develop a standard curve.
Acknowledgements

I would like to first thank my industry collaborators for all of their time and effort in making this project a success. I am also extremely grateful to my major professor, Dr. Powell, for pushing me to become better in every aspect of my work and especially for the continuum of help with my writing.

My deepest gratitude to all the friends, family, professors, and colleagues that encouraged my to keep going and a special aside to those that helped me with the daunting task of proofreading.

Sarah M. Reehl
Contents

Abstract ........................................................................................................... iii
Public Abstract .............................................................................................. v
Acknowledgements ......................................................................................... viii
List of Figures ................................................................................................. xi
List of Tables ................................................................................................... xiii
1 Introduction .................................................................................................. 1
  1.1 Enzyme Linked Immunosorbent Assay (ELISA) ....................................... 1
     1.1.1 History ............................................................................................. 2
     1.1.2 ELISA Basics .................................................................................. 2
     1.1.3 The Standard Curve ....................................................................... 3
  1.2 Introduction to Likelihood and Current ELISA Industry Methods .......... 4
  1.3 Current Methods ..................................................................................... 6
     1.3.1 Weighting Sum of Squares and Likelihood ...................................... 6
     1.3.2 Data Transformations ..................................................................... 7
  1.4 Costly Problems Unique to Quantitative ELISA .................................... 7
     1.4.1 Outlier Detection ........................................................................... 8
     1.4.2 Parameterization .......................................................................... 8
     1.4.3 Standard Wells ............................................................................. 8
  1.5 A Bayesian Approach ............................................................................. 9
     1.5.1 Thomas Bayes and Posterior Probability ...................................... 9
     1.5.2 Industry Applications of Bayesian Perspective .............................. 10
  1.6 Goals ...................................................................................................... 11
2 Methods ...................................................................................................... 12
  2.1 Collecting Prior Information .................................................................. 12
     2.1.1 Setting the Priors ......................................................................... 12
     2.1.2 Posterior Probability .................................................................... 13
     2.1.3 Quality of Fit Measures Defined .................................................. 15
  2.2 Developing Outlier Detection .................................................................. 16
  2.3 Developing a Weighting of Prior Information ....................................... 21
  2.4 Developing a Standard Design Using Fewer Sample Wells .................. 24
3 Results ......................................................................................................... 26
  3.1 Outlier Detection ................................................................................... 26
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2</td>
<td>Weighting Prior Information</td>
<td>28</td>
</tr>
<tr>
<td>3.3</td>
<td>Standard Design with Fewer Sample Wells</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>Discussion and Conclusion</td>
<td>35</td>
</tr>
<tr>
<td>4.1</td>
<td>Summary</td>
<td>35</td>
</tr>
<tr>
<td>4.2</td>
<td>An Example Application</td>
<td>37</td>
</tr>
<tr>
<td>4.3</td>
<td>Future Work</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Bibliography</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>Appendices</td>
<td>41</td>
</tr>
</tbody>
</table>
### List of Figures

1.1 Immunometric immunoassay that functions using very specific antibody antigen reactions [21]. ................................................................. 3

1.2 Summary of the effects of parameters on the five parameter logistic curve common to ELISA practice. .................................................. 4

2.1 Histograms of estimates for each of the five parameters using Analyte 1 data. These histograms may help in assuming a distribution on each parameter. ................................................................. 14

2.2 Boxplot representations of each standard dilution pixel intensity from a family of fits for root transformed Analyte 1. The whiskers shown mark boundaries for $\alpha = 0.05$ and the box edges represent $\alpha = 0.25$. ............. 17

2.3 Illustration of outlier detection for Analyte 1 at $x_8$. .......................... 18

2.4 An example of one replicates series of perturbed Analyte 1 data. The perturbations occur in the first and fourth concentrations. ............. 19

2.5 The changes in $R(\alpha)$ according to $\alpha$ in perturbed Analyte 1 data. Contrasts between the nominal case and the removed case, measured by the function $R(\alpha)$, show an $\alpha$ near 0.05 as optimal as the tangent line at $\alpha \approx 0.05$ has the largest slope. .................................................. 20

3.1 Percentage of data points with a backfit between 0.80 and 1.20 both before and after outlier detection and removal. We see no change in these percentages as no points were removed at this level. .............. 26

3.2 Percentage of data points with a backfit between 0.80 and 1.20 both before and after outlier detection and removal. Several replicates contained one to two detected outliers, which were removed. ............................. 27

3.3 Percentage of data points with a backfit between 0.80 and 1.20 both before and after outlier detection and removal. The percentages improve due to the loss of several data points in each replicate. ......................... 27

3.4 Percentage of data points with a backfit between 0.80 and 1.20 in series eight of Analyte 4. The initial percentage is already at a maximum when $\alpha = 0$ and when $\alpha = 0.05$ the percentage degrades on remaining data points. ................................................................. 28

3.5 Density plot of averaged Analyte 1 SSE, from curves fitted using a wSSE, that illustrate the distance of the curve from the data when fitting data with two removed replicate pairs. ................................. 30
Parameterized curves for each of the three designs using fewer sample wells for the standard where the removed \( x_j \) are shown as ‘O’ and the remaining \( x_j \) used for parameterization are shown as ‘X’. Each curve is fit with \( m = 0 \) (black) and \( m = m_{\text{optimum}} \) (green). The difference of fit in the ten point design is nearly undetectable, with a small change in the fit near \( x_8 \) increasing the LLOQ. The eight point design does not benefit from \( m \) in this series. The six point design shows a noticeable difference in fit when using \( m \), where \( m = m_{\text{optimum}} \) again improves the LLOQ.

A simulated missed wash replicate where the first attempt at fitting this standard with wNLL provides two of eight points in an acceptable backfit. Outliers are detected in the first and second concentrations, then the fit improves to acceptable backfits in five of six remaining points. We capture all six data points to an acceptable backfit by tuning the weight on prior information.
List of Tables

3.1 Approximations on $\Delta m$ by analyte according to a nullcline analysis from 2.7 in Section 2.3. These approximations are averaged over thirteen replicates for each analyte. .......................................................... 28
3.2 Percentage of data in an acceptable backfit for the nominal MLE and the nominal MPLE curve fitting procedures. Also shown are the minimum and maximum $m$ that improved acceptable backfit percentages in some replicate series by analyte and our nullcline estimate of $\Delta m$. ......................... 29
3.3 The percentage of replicate series with an LLOQ at the lowest concentration. Percentages are out of thirteen replicate series for each unperturbed analyte. ................................................................. 30
3.4 New standard designs with three, four, and five replicate point removals. The removed replicates are indicated by 'OO'. The replicates that remain and parameterize the model are indicated by 'XX'. ................................. 31
3.5 Percentage of data in an acceptable backfit for the nominal design with all seven dilutions and each of the reduced dilution schemes. Percentages are out of sixteen data points averaged across thirteen replicates for each of the six analytes. ......................................................... 31
3.6 The minimum and maximum $m$ used in testing by analyte. A replicate series may have an optimal fit with $m = 0$ or an $m$ between the minimum and maximum shown. ......................................................... 32
3.7 Percentage of replicates with an LLOQ at the lowest initially measured concentration summarized by analyte. Percentages are shown for both MLE and MPLE across each of three, four, and five pair replicate point removals. ......................................................... 32
Chapter 1
Introduction

Since Baron Kitasato Shibasaburō began studying antibodies in 1890, the detection of antigens in serum uses the “lock and key” properties between antigens and antibodies [1]. Differing segments of DNA drive the mechanisms which bind antigens and antibodies under both chemically attractive forces and complementary shapes, emulating a “magnetic lock and key” [15]. When antibodies and antigens bind together, a number of detectable chemical reactions occur which then acts as an indicator for the presence of the antibody-specific antigen in a solution [15, 18]. As antibodies are extremely specific in their binding properties, results from serum tests naturally carry a high degree of certainty. To enhance this qualitative certainty, studies emerged seeking a method for qualitative results that would assess not only the presence, but the amount, of antigen in a solution. Arriving in 1971, the enzyme linked immunosorbent assay revolutionized modern medicine with the ability to produce these quantitative results [4].

1.1 Enzyme Linked Immunosorbent Assay (ELISA)

Assays are the analytic procedures used to measure target entities, called analytes, in a sample. They are commonly used for indicative testing in contemporary medicine [22]. The enzyme-linked immunosorbent assay (ELISA) specializes in detecting and quantifying trace amounts of an analyte in a sample using the same properties of antigen-antibody reactions that Baron Shibasaburō had developed nearly 100 years prior [9]. However, the subtle difference that allows ELISA to measure analyte concentrations is that it relies on an enzyme label to generate a quantifiable signal from the in vitro antibody reactions. We turn to Scandanavia for the origins of this remarkable invention [9].
1.1.1 History

In the 1970’s, two groups of Swedish and Finnish scientists independently and simultaneously developed this special type of enzyme labeled immunoassay. Swedish scientists, Peter Perlmann and Eva Engvall, published their first paper using quantitative enzyme immunoassay measurements on a serum of rabbit protein [4]. Soon after, the Finns Weemen and Schuurs published on the human chronic gonadotropin concentrations independently using the same method. Several new diagnostic tests grew out of these preliminary findings, including tests for hepatitis B, rubella, and a system for HIV detection. The practice became clinical in 1980 and has since revolutionized the analytical investigation process, especially in the advancements of HIV studies [9].

1.1.2 ELISA Basics

Say that a clinical physician is testing a small quantity of serum for HIV. In this example, the serum is the sample and HIV is the analyte. The physician generates antibodies for the analyte by a process of vaccination, in which dead pieces of the virus are used to build an immune response, creating the antibodies associated with that analyte. Using the interaction between the antibody and the analyte, a cascade of reactions generates a detectable signal. At this point, the power of ELISA becomes apparent due to the extreme specificity of the antibody to its target analyte, the strength with which the antibody binds to the target, and the wide range of analytes that can produce antibodies for testing [21].

The first step to this assay is to bind the manufactured antibody onto a plastic surface, normally called a plate (shown in Figure 1.1). The sample and the plate come into contact and the HIV-antibody dwelling on the plate binds to the target analyte in the sample. The plate then receives a wash with a solution of new HIV-antibodies that are labelled with a radioactive isotope. These labelled antibodies bind to the analyte already bound by the plated antibody forming a “sandwich.” This binding process activates the label, which releases a quantum level of light in the case of ELISA. After incubation and a series of washes, a reader captures the amount of light generated by the reaction [21]. Larger quantities of an analyte present in a sample will yield brighter pixel intensities to the reader.

Although some immunoassays, such as a publicly available pregnancy test, only give qualitative results [21], we are specifically interested in immunoassays that produce
quantitative results. In partnership with a biotech-company, we explore mechanisms and issues unique to quantitative ELISA. As the design of these types of assays vary slightly from test to test, and even from one reading machine to the next, each analyte and plate must be calibrated according to a standard. A standard curve is the curve which results from parameterizing a model of pixel intensities in response to inputs of known concentration. The known concentrations are serial dilutions of a sample so that the curve can be calibrated across an entire spectrum of possible analyte concentrations [3]. Given a standard curve, a sample of unknown concentration can be quantified using its observed pixel intensity, working backwards to locate the model concentration that coincides with the the observations [3].

1.1.3 The Standard Curve

ELISA is a widely accepted practice among users. The industry has converged on a five parameter logistic standard curve [3, 5]:

\[ Y(x) = d + \frac{a - d}{1 + \left( \frac{x}{c} \right)^b} , \]

1. Figure 1.1: Immunometric immunoassay that functions using very specific antibody antigen reactions [21].
Figure 1.2: Summary of the effects of parameters on the five parameter logistic curve common to ELISA practice.

where:

\[ Y(x) = \text{pixel intensity or response to analyte concentration, } x \]

- \( a \) = lower limit of quantification
- \( d \) = upper limit of quantification
- \( c \) = expected concentration halfway between \( a \) and \( d \)
- \( b \) = slope factor related to the change of response with increasing concentration
- \( g \) = asymmetry factor

Figure 1.2 is a summary of the effects of the parameters on the five parameter logistic curve. Normally, samples are run in replicate on a seven part serial dilution where the sample is diluted the same amount in each step. The standard then consists of a sample of known concentrations, diluted in serial, so that the signal response can be measured for each series of known concentrations. Each dilution series is also run in replicate to catch any aberrations within the standard. The log of known concentrations, \( x_j \), and the resultant pixel intensities, \( y_j \), are used to determine the parameters in Equation 1.1. To quantify an unknown sample, we invert the standard curve to estimate an unknown concentration using the measured signal response. ELISA’s predictive abilities are completely determined by the accuracy of the curve fit to the standard.

1.2 Introduction to Likelihood and Current ELISA Industry Methods

Fitting curves to data has been no great mystery, at least since the 18th Century when Carl Gauss introduced the notion of minimal sum squared error [17]. R.A. Fisher
later systematized the minimal error approach and introduced the notion of likelihood [2], which we will review in this section. Let $y_j$ denote the observed pixel intensity at concentration $x_j$ and let $Y_j$ denote the five parameter logistic model-predicted value from concentration $x_j$. The $y_j$ are not perfectly predicted by the model so

$$y_j = Y_j + \epsilon_j,$$

where $\epsilon_j$ is assumed to be a random error. In fitting a curve to this data, we aim to select parameters that minimize the difference between our observations and model predictions, that is to minimize independent and identically distributed (iid) $\epsilon_j$ [6], with a mean of zero. We can also view this as parameterizing our model to maximize the probability of observing the data given our model. The probability of observing a normally distributed, iid error follows the normal probability density function:

$$P(\epsilon_j|\text{parameters}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(\epsilon_j)^2}{2\sigma^2}\right). \quad (1.2)$$

We can modify Equation 1.2 according to the definition of $\epsilon_j$ in order to see that:

$$P(y_j|\text{parameters}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(y_j - Y_j)^2}{2\sigma^2}\right) \quad (1.3)$$

where we call $P(y_j|\text{parameters})$ the “likelihood” of the $j$th observation. Assuming observations are independent, we can multiply the probabilities of each observation to obtain the likelihood of all observations given our model:

$$L(\text{data}|\text{parameters}) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(y_j - Y_j)^2}{2\sigma^2}\right). \quad (1.4)$$

The most suitable parameters would maximize Equation 1.4. This maximization technique is commonly called maximum likelihood estimation (MLE). To make MLE easier computationally, we minimize the negative log of the likelihood (NLL)

$$\text{NLL} = -\log(L) = \sum_{j=1}^{n} \left[ \frac{(y_j - Y_j)^2}{2\sigma^2} + \frac{1}{2} \log(2\pi\sigma^2) \right]. \quad (1.5)$$
It is easy to show that MLE with normal errors is equivalent to minimizing a sum of squares error (SSE)

\[
\text{SSE} = \sum_{j=1}^{n} (y_j - Y_j)^2 , \tag{1.6}
\]

where \( n \) is the total number of data points \([12]\). In engineering settings, minimum SSE is used habitually, without reference to the likelihood function in the background.

### 1.3 Current Methods

At present, the five parameter logistic model is the accepted model for ELISA data. Equation 1.1 is a smooth, continuous curve that seems to appropriately capture elements of ELISA chemical response \([6]\). We cannot change standard guidelines that use the four or five parameter logistic model, as these guidelines have been in place since the clinical development of ELISAs \([14]\). As a result, we cannot improve the model fit by changing the model. Instead, only changing the parameterization of the model can improve the fit. Insights to the problems facing ELISA manufacturers and users were given to us by the company that provided the data.

In addition to rigid guidelines, there are procedures in place unique to ELISA sensitivity. Frequently, ELISA is used to detect very small amounts of analyte in a sample, such as searching for traces of nut protein on commercial food processors \([20]\). Consequently, this means that the curve fitting routine needs to be particularly sensitive in the lower concentration spectrum. If a nut protein standard showed large changes in replicates of \( y_j \) at the high concentration \( x_j \), then we are likely to lose the most important information about the replicates of \( y_j \) at low concentrations, where the bulk of our test samples occur. At present, industry practice includes both data transformations and weighting mechanisms to improve fits in the low concentration spectrum.

#### 1.3.1 Weighting Sum of Squares and Likelihood

In order to increase sensitivity at low concentrations, each \( \epsilon_j \) from Equation 1.6 is often multiplied by the inverse variance of the responses for the \( j \)th concentration \([5]\):

\[
\text{wSSE} = \sum_{j=1}^{n} c * \frac{(y_j - Y_j)^2}{w_j} \tag{1.7}
\]

where:

\[
w_j = (y_{j,1} - y_{j,2})^2 \tag{1.8}
\]
and
\[ c = \frac{1}{\sum_{j=1}^{n} \frac{1}{w_j}} \] (1.9)

so that \( \sum_{j=1}^{n} \frac{c}{w_j} = 1 \). The corresponding weighted NLL is

\[
wNLL = \sum_{j=1}^{n} \left[ \frac{c(y_j - Y_j)^2}{w_j 2\sigma^2} + \frac{1}{2} \log(2\pi\sigma^2) \right]
\] (1.10)

This particular weighting scheme ensures that the standard responses with the smallest variance between replicates, generally low concentrations, will be given more weight than the standard responses with the largest variance, which also happens to be those with the largest response [5].

### 1.3.2 Data Transformations

Transformations may potentially standardize the random error associated with each data point. Most common is the root transformation, which inputs the square root of observed data so that the model actually fit is a root of Equation 1.1,

\[
\sqrt{Y(x)} = \sqrt{d + \frac{a - d}{\left(1 + (\frac{x}{c})^b\right)^g}}.
\] (1.11)

With the combination of weighting and data transformations, we can ensure sensitivity at the lower concentrations. In this study, we transform every observation under a root transformation for every analyte.

### 1.4 Costly Problems Unique to Quantitative ELISA

In the context of industry, ELISA materials are manufactured by companies that provide analyte-specific 96-well plates to consumers. Each well is expensive for the client to load, with an average development cost of $30,000 - $150,000 per analyte-specific assay and an average continued production cost of $50 - $350 per 96-well plate [8]. While each well carries a price, the plate uses 12-16 wells to develop the standard and only the remaining 80-84 wells are dedicated to testing samples of unknown concentration [14]. To maximize the number of unknown samples on a plate, the standard replicate series exists as a minimum of only two replicates at each concentration. Already at a minimum of two replicates per plate, the standard still uses 15% of available testing space.
In addition, problems occur within ELISA data collection, dilutions being improperly
diluted, mixed or washed, along with inappropriate incubation times. Regardless of the
source of error, bad data occurs and affects the entire assay, especially the parameter-
ization of the standard. If the standard cannot properly approximate \( y_j \) of known \( x_j \),
then the inversion of the standard will not accurately quantify samples of unknown \( x_j \).
When samples are not accurately quantified, the assay becomes useless and costly to the
consumer.

1.4.1 Outlier Detection

Any out of place, or outlying data will likely result in poor parameterization of
Equation 1.1. As the standard exists as a series of only two replicates it impossible to
quantify and neglect outliers at any particular sample concentration where at least two
replicates are required for parameterization. Without the ability to incorporate a third
replicate of the standard, the entire plate is lost due to only one or two bad data points.
We aim to identify and remove outlying data to maintain an accurate parameterization
of the standard curve.

1.4.2 Parameterization

The consumer faces a cost, even if the data collected is non-anomalous, when the
likelihood chooses parameters from a local minimum error rather than a global minimum
error. As this parameterization problem is nonlinear, convergence on an appropriate
minimum causes problems numerically. To avoid this type of cost, we want to make the
curve fitting robust and assess the appropriate minima in parameter selection.

1.4.3 Standard Wells

Not only does bad data and poor parameterization add cost to the consumer, but
the wells dedicated to setting the replicate standard are wells lost to sample tests. If
the dilution scheme for setting a standard used only ten, eight, or even six wells, the
consumer would free up two to six wells for testing samples of unknown \( x_j \). In order to
keep test economically feasible, we also want to gather as much information as we can
about a sample standard using as few data points as possible.
1.5 A Bayesian Approach

Most companies run several standards in the process of developing an ELISA for a particular analyte. Our idea is to utilize this prior information in a Bayesian approach to influence the probability of parameters in fitting the five parameter standard curve. We cannot entirely replace the standard, as each plate differs slightly from the last, but we may be able to use past parameterizations on an analyte to robustly guide present parameterization with missing or generally fewer data points. Using prior information to guide current inference is the goal of Bayesian statistics, in which the probability of the model is adjusted by previously collected data [7].

1.5.1 Thomas Bayes and Posterior Probability

Reverend Thomas Bayes recognized a distribution on parameters as a prior distribution since it refers to the distribution of a “degree of belief” on the parameter before any data has been taken. The quantitative recipe for including prior information is Bayes’ rule, a relation between conditional probabilities on two events. Bayes’ rule states that [16]:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (1.12)$$

where:

- $P(A|B)$ = the posterior probability for A given B
- $P(A)$ = the probability of A prior to information about B
- $P(B)$ = the probability of B prior to information about A
- $P(B|A)$ = the likelihood function for B given A

After data collection occurs, especially collection that holds weight on our parameters, we update the prior distribution of parameters to create a posterior distribution, having observed more information from the outcomes of data collection. S. J. Press calls it “...a formal procedure for merging knowledge obtained from experience, or theoretical understanding of a random process, with observational data” [16]. We intend to use the
posterior distribution in modelling ELISA data where [7]:

\[
P(\text{Parameters}|\text{data}) = \text{the posterior probability} \\
\]
\[
P(\text{data}|\text{Parameters}) = \text{the likelihood of the model} \\
\]
\[
P(\text{Parameters}) = \text{the “prior” distribution of parameters} \\
\]
\[
P(\text{data}) = \text{a normalization constant} \\
\]

As such, \( P(\text{data}) \) is a constant across parameters and doesn’t affect parameterization. So, the posterior probability, \( P(\text{Parameters}|\text{data}) \), is

\[
P(\text{Parameters}|\text{data}) = P(\text{data}|\text{Parameters}) \times P(\text{Parameters})
\]

where \( P(\text{data}|\text{Parameters}) \) is the likelihood of the parameters, which is the same likelihood from Section 1.2. We will fit all of the pre-existing data provided and use each parameterization to form the prior distribution, \( P(\text{Parameters}) \), on each of the five parameters. Then, the posterior probability for a given set of data is

\[
P(\text{Parameters}|\text{data}) = \mathcal{L}(\text{data}|\text{Parameters}) \times P(\text{Parameters}) \tag{1.13}
\]

that is, the product of the likelihood of the parameters given the data and the probability of the parameters given prior information. The best parameters are selected using maximum posterior likelihood estimation, or MPLE, that is a product of prior parameter probability and normal likelihood. If we have a collection of data for an analyte, we should be able to use that prior information to help guide our curve fitting algorithm toward likely parameter values. In other words, we are looking to maximize the probability of our parameters, given the data, according to previous data/parameter interactions.

### 1.5.2 Industry Applications of Bayesian Perspective

Industry is no stranger to Bayesian approaches. The pharmaceutical industry uses previous outcomes between phases of pharmaceutical drug development to predict the probability of success in the final phase [19]. In ecology, prior information helps guide population studies with few observations as well as providing guidance for distinguishing
between competing hypotheses [11]. We plan to utilize data generated in the development and quality control phases of the immunoassy in order to influence maximum likelihood estimation. In a deviation from true Bayesian statistics, our methods do not include new information about the posterior distribution of our parameters. We are more interested in controlling the incremental search for parameters based on parameter distributions from lab-derived prior parameters [7].

1.6 Goals

Our idea is to address outliers, parameterization improvements, and the design of standard dilutions using a Bayesian perspective to introduce prior information. First, we will build prior distributions by fitting the five-parameter curve to “good” data using wSSE. Distributions of credible responses at particular concentrations will be developed by sampling the priors and evaluating the five parameter curve at those samples. Outliers can then be characterized in an analysis of the tails of the posterior response distribution. An appropriate credible level will be developed for a particular analyte and then tested against others. In parameterization, we will avoid spurious local error minima by maximizing the posterior likelihood for each analyte. A sensitivity analysis will determine a weighting for prior parameter distributions that maintain constant information in the face of new data.

We then remodel the standard design, utilizing fewer standard sample concentrations. We also test alternate standard designs by removing dilution concentrations sequentially from the weighted likelihood and testing a coefficient of determination using all available data. The design which cause the least degradation to the coefficient of determination across all replicate series of all analytes is the proposed design. For removals of three to five dilutions we assess the performance of the standard parameterization by the percentage of acceptably fit data and the percentage of replicate series with acceptably fit data in the low concentration spectrum. We show assessments for parameterization both with and without prior information. Finally, we comment on further development of the priors, problems in weighting and transforming data, and predicting unknown concentrations without using a reference model.
Chapter 2
Methods

2.1 Collecting Prior Information

Our data consists of six different analytes run in sets of thirteen replicate series of a seven part serial dilution. This data set incorporates good samples that produce accurate standard curves as well as problematic data that results in less accurate parameterization of the standard. The abbreviated data appear in Appendix A. We then use the techniques described below to develop prior distributions for each analyte.

2.1.1 Setting the Priors

Quality control procedures throughout all stages of product development incidentally produce several replicates of a standard for any of the analytes in development. Normally, the data feeds into a curve fitting routine that numerically approximates a minimum wSSE, with respect to each parameter, and a global error minimum under the combination of parameters. We utilize MATLAB’s fminsearch function, which employs the Nelder-Mead simplex algorithm to shrink a simplex around the initial parameter guesses until the diameter of the simplex is less than a specified tolerance, in order to numerically approximate each parameter [10]. In addition, our methods use both likelihood and prior information to parameterize the curve rather than wSSE. See Appendix C.1 for specific MATLAB parameterization code.

Let \( \theta_{kp} = [\theta_1 \ \theta_2 \ \theta_3 \ \theta_4 \ \theta_5] \) where \( \theta_1 = a, \ \theta_2 = b, \ \theta_3 = c, \ \theta_4 = d, \) and \( \theta_5 = g. \) First parameterizing all of the quality control replicate series and keeping the best-fit instances, we develop a prior collection of each of the five parameters. We then use these parameters to develop empirical prior parameter distributions with mean \( \theta^*_{kp} \) and variance \( \nu_{\theta_{kp}}. \) If we parameterize a single replicate series with estimated parameters \( \theta_{kp}, \) these estimates will have a random iid error associated with fitting noisy, imperfect data.
\[ \theta_{kp} = \theta_{kp}^* + \epsilon_{kp}, \]

where \( \theta_{kp}^* \) is the mean of the \( kp \)th prior. Assuming \( \epsilon_{kp} \) follows a normal distribution with a zero mean

\[ P(\theta_{kp}) = \frac{1}{\sqrt{2\pi\nu_{kp}^2}} \exp \left( \frac{-(\theta_{kp}^2 - \theta_{kp}^*)^2}{2\nu_{kp}^2} \right), \]

where \( \nu_{kp} \) is our best estimate on the variance of the parameters. For all analytes we develop the prior distributions by fitting root transformed data, keeping only the parameterization instances with acceptable quality of fit measures defined below. For example, using Analyte 1 data and five instances of acceptable parameterization, the prior distributions are shown in Figure 2.1 histograms seen in Figure 2.1. In our methods, we assume a normal distribution on the parameters as the amount of data we have to make such a judgement is low. See Appendix B for a table of estimates on prior distributions.

### 2.1.2 Posterior Probability

We incorporate these developed prior distributions into a posterior distribution. A new parameter \( m \), which we will interpret as a weight of prior data, controls for the amount of prior information we will incorporate into the posterior. If new information is only slightly similar to prior information, we want to use less of the prior information to influence parameterization. Likewise, if the new and prior information are very similar, parameterization would benefit from the use of more prior information. Then, using the posterior probability from Bayes’ Theorem as a product of normal likelihood and prior probability

\[ P(\text{Parameters}|\text{data}) = \frac{1}{P(\text{data})} \mathcal{L}(\text{data}|\text{Parameters}) \ast (P(\text{Parameters}))^m, \]

where we assume \( \frac{1}{P(\text{data})} = 1 \), a constant across all parameters, so that the posterior distribution follows

\[ P(\text{Parameters}|\text{data}) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left( \frac{-c(y_j - Y_j)^2}{w_j^2\sigma^2} \right) \ast \prod_{kp=1}^{5} (P(\theta_{kp}))^m. \]
In a similar fashion to minimizing a wNLL, we minimize a negative log posterior (NLP) in a technique of maximum posterior likelihood estimation (MPLE):

$$\text{NLP} = \text{wNLL} + m \sum_{k_p=1}^{5} \frac{(\theta_{k_p}^* - \theta_{k_p})^2}{2\nu_{k_p}^2},$$

(2.1)

where $m$ is a weighting parameter of the prior information. The wNLL from Equation 1.10 now utilizes information from previous parameterization instances in the NLP. In effect, the priors cause the model to select parameters that are similar to past parameterization instances over parameters that are grossly similar. We hope that this
feature will aid in avoiding the spurious minima that occur with complex minimization problems, such as minimization across five parameters with very few data points.

2.1.3 Quality of Fit Measures Defined

The accuracy of ELISA quantification relies on how closely the standard curve approximates real data. Quality of fit measures allow the user to judge how well parametrized Equation 1.1 fits the data and allows for a metric to compare different fits.

Backfit

Backfit is a measure of how well the observed data are predicted by the fitted five parameter curve. We can use a known intensity and divide it by the predicted intensity for a ratio of the difference.

\[
\text{backfit}_j = \frac{y_j}{Y_j}
\]

(2.2)

Industry standard requires the backfit value to lie between 0.80 and 1.20 for the fit to be considered acceptable at a given concentration[14]. We also use backfit to define the data we use for our prior information as parameterization instances that backfit all 16 data points between 0.80 and 1.20.

Coefficient of Determination (CD)

The coefficient of determination (CD) is a more continuous way to measure the overall fit of the model to the data. For data under a root transformation:

\[
\text{CD} = 1 - \frac{\sum_{j=1}^{n} (\sqrt{y_j} - \sqrt{Y_j})^2}{\sum_{j=1}^{n} (\sqrt{y_j} - \bar{\sqrt{Y}})^2} = 1 - \frac{\text{SSE}}{\text{MSE}}
\]

(2.3)

The CD uses a ratio of SSE from Equation 1.6 and the mean square error (MSE) in order to calculate the amount of variance explained by the model[13]. A CD \(\approx 1\) indicates that most of the variance is explained by the model as the ratio of errors is small. Generally, as CD increases toward one, the number of acceptable points increases.

Lower Limit of Quantification (LLOQ)

The most valuable part of the curve lies in the low concentrations. The LLOQ determines the lowest concentration of an analyte sample that can be quantified with
acceptable precision and accuracy[14]. As data points tend to fit worse in the lowest concentration range, the LLOQ is measured from the highest concentration to lowest concentration; LLOQ is the data point previous to the first data point with an unacceptable backfit. Every lower concentration point beyond the first unacceptable backfit in the lower concentrations is then excluded from the range of quantification[14].

2.2 Developing Outlier Detection

With only a small population of data points to determine parameter values, the data must be as closely representative of the empirical curve as possible. Any data that is dissimilar, or outlying, to the rest of the data set can potentially disrupt parameterization. Traditionally, outlying data points are identified using statistics on a population sample, representing a larger, whole population. ELISA has only two replicates at each concentration, so that detecting outliers with traditional statistics simply doesn’t make sense. However, armed with prior information, we can identify outliers by simulating a much larger sample population.

Recall that $P_{kp} (\theta_{kp})$ is a prior parameter distribution with mean $\theta_{kp}^*$ and variance $\nu_{\theta_{kp}}$. We form a “family of fits” $\mathcal{Y}$, developed by sampling $\theta_{kp}$ from each $P_{kp} (\theta_{kp})$. Let $\hat{\theta}_{kp,i}$ be the $i$th set of parameters sampled at random from $P_{kp} (\theta_{kp})$, where caret denotes a random sample. We then simulate $Y_i(x_j) = Y(x_j, \hat{\theta}_{kp,i})$ as a response pixel intensity of dilution concentration $x_j$, based on the sampling of priors. These simulations give us a collection of response pixel intensities

$$\mathcal{Y}(x_j) = \{Y(x_j, \hat{\theta}_{kp,i})\}_{i=1}^I$$

that now form a population of size $I$ at concentration $x_j$. We bin the simulated observations into $s$ different bins and let the frequency of each observation, $z_i$ for a cumulative sum of the bins normalized by the total number of simulated observations at $x_j$. If $z_i$ lies outside of a threshold, $\alpha$, for the $x_j$th concentration response in $\mathcal{Y}$, we can designate the response as an outlier and remove it from the curve fitting procedure. An observed pixel intensity then corresponds to some $z_i$ at concentration $x_j$ and can be identified as an outlier if the frequency is not contained between $z_i \in [\alpha, 1 - \alpha]$ as seen in Figure 2.2(c). Since bin centers are unlikely at arbitrary $\alpha$, we use linear interpolation between the last bin below and the first bin above to accurately estimate the threshold. See Appendix C.2 for specific MATLAB outlier detection code.
Using Analyte 1 to illustrate the process, we first gather our priors using replicate series that fit all 16 data points to a backfit value between 0.80 and 1.20. Using parameter means and standard deviations, we form a distribution on each of the parameters. Randomly selecting combinations of parameters from these distributions forms the family of fits seen in Figure 2.2(a). Using the family of fits, we then form distributions on each of the pixel intensities for the current standard dilution scheme. The distribution of $\mathcal{Y}(x_8)$ is shown in Figure 2.2(b), which demonstrates a population of responses we might expect to see at this concentration based on prior information. The boundaries shown mark the pixel intensities designated as outliers if $\alpha = 0.05$. The outlier pixel intensities shown in Figure 2.2(b) are determined using pixel intensities such that the frequency of the intensity is $z_i < 0.05$ and $z_i > 0.95$ shown on the cumulant in Figure 2.2(c).

Extending the process to each $x_j$ for root transformed Analyte 1, we specify a threshold that will identify any $y_j$ as an outlier if it falls beyond the constructed boundary. A box plot of each of the resultant Analyte 1 $\mathcal{Y}(x_j)$ for the standard dilution scheme is shown in Figure 2.3.

![Distribution of Pixel Intensities for Root Transformed Analyte 1 Family of Fits [$\alpha = 0.05$]](image)

**Figure 2.3:** Boxplot representations of each standard dilution pixel intensity from a family of fits for root transformed Analyte 1. The whiskers shown mark boundaries for $\alpha = 0.05$ and the box edges represent $\alpha = 0.25$.

We can set any $\alpha$, to label $y_j$ as outliers if they occur at an intensity outside of the frequency boundary. The whiskers in Figure 2.3 show the boundaries for $\alpha = 0.05$. Increasing $\alpha$ shrinks the whisker length and the algorithm becomes more exclusive. Depending on the $\alpha$, data points are identified as outliers or non-outliers and then
(a) A family of 1000 different fits developed by random parameter combinations collected from the prior distributions of $a$, $b$, $c$, $d$, and $g$ in the five parameter curve.

(b) A histogram of simulated pixel intensities at high concentration for root transformed Analyte 1. We set $\alpha = 0.05$ and exclude pixel intensities outside of the shown markers corresponding to $\alpha$ and $1 - \alpha$.

(c) A cumulant of simulated pixel intensities at high concentration for root transformed Analyte 1. Any pixel intensity with frequency $z_i < 0.05$ and $z_i > 0.95$ is defined as an outlier.

Figure 2.2: Illustration of outlier detection for Analyte 1 at $x_8$. 
removed or selected accordingly for the fitting procedure. We would wish to use some $\alpha$ that eliminates the fewest number of data and improves poor quality of fit measurements for given replicates of any analyte. Ultimately, choosing $\alpha$ is under the user’s control.

To test a reasonable $\alpha$, we define a range for $\alpha$ from 0.01 to 0.25 and set out to detect and remove outliers. Using a replicate Analyte 1 data, we perturb replicates in the first and fourth concentrations, reducing one replicate each by a random factor near 100 units of intensity. The example replicate series shown in Figure 2.4 is one such perturbation. We degrade the data at these concentrations to deliberately influence the parameterization of these replicates. Then, we can use the poorly fitted “bad” data to determine an $\alpha$ that restores acceptable backfits to the unperturbed $y_j$. Let the measure of goodness of fit, $mog(y_j, \alpha)$, be the percentage of non-outlier data points with an acceptable backfit

$$mog(y_j, \alpha) = \frac{n - n_b}{n}$$

where $n_b$ denotes the number of observations, $y_j$, with $y_j \in (\text{backfit}_j \leq 0.80) \cap (\text{backfit}_j \geq 1.20)$. For example, a perfect fit has all data points, $y_j$, in an acceptable backfit with no outliers detected so that $mog(y_j, 0) = 1$. In contrasting the data before outlier removal, or the nominal case, and after outlier removal, the removed case, we can let a function, $R$, measure improvement as the difference in goodness of fit between the nominal case and the removed case, normalized by the difference in goodness of fit between the nominal case and the perfect case:

$$R(\alpha) = \frac{mog(y_j, \alpha) - mog(y_j, 0)}{(1 - mog(y_j, 0))}.$$
Removing data will eventually lead to diminishing returns in $R(\alpha)$ where the benefits per increase in $\alpha$ become smaller. This implies that $\alpha$ is constrained by some optimum where increases in $R(\alpha)$ are the largest with the least amount of data removed, or the secant with the largest slope that tends to the tangent line at $R(\alpha)$ so that

$$\alpha_{\text{optimum}} = \max \left( \frac{R(\alpha)}{\alpha} \right).$$

At some point, a large proportion of the data is removed and the backfit percentage will become artificially inflated, to the point of perfection. Iterating through each $\alpha$, we began looking for the optimum $\alpha$, where the largest increase in $R(\alpha)$ occurs for the smallest increase in $\alpha$. Removing data beyond the optimal $\alpha$ provides no real benefit to MPLE as the improvements in $R$ occur due to artificial inflation. Figure 2.5 shows that initially, $\alpha \approx 0.01$ does not improve the goodness of fit as this threshold was not high enough to remove both of the obvious outliers at the same time. However, an $\alpha$ near 0.05 provides a 60% improvement over the nominal case. The tangent line at $\alpha \approx 0.05$ clearly has the largest slope designating this $\alpha$ as optimal. Removing points after $\alpha \approx 0.05$, in this extreme data set, damages $R(\alpha)$ until we reach artificial inflation near $\alpha \approx 0.20$. The appropriate $\alpha$ will improve the backfit percentage while removing the least amount of data, that is the appropriate $\alpha$ increases $R(\alpha)$ with the smallest change in $\alpha$ as seen in Figure 2.5.
To test the method on other analytes, we set $\alpha = 0.01$, $\alpha = 0.05$, and $\alpha = 0.10$ to test outlier detection on replicate series of all analytes. For each set of replicates we measure the percentage of data in the 0.80 to 1.20 backfit range after outlier detection and removal. We test the quality of each $\alpha$ using the percentage of remaining data points in an acceptable backfit while monitoring the volume of data removed. We selected Analyte 1 to initially test all of our methods and tested the matured methods on the remaining analytes in our data set. We used good data, where all $y_j$ have a backfit value between 0.80 and 1.20, as prior information in developing prior probabilities on the parameters. We then extend each developed method to the other analytes in our data set.

2.3 Developing a Weighting of Prior Information

If an abundance of prior information exists for an analyte, our confidence raises in using the priors in MPLE. The parameter $m$ in Equation 2.1 controls the certainty we have on prior information by weighting the likelihood estimation with respect to the priors. If $m = 0$, effectively no prior information is utilized and the fitting procedure only relies on the likelihood. As $m$ increases, MPLE begins to incorporate the prior probability associated with each parameter. If the likelihood starts to stray towards a spurious minimum, the prior parameters would discourage selecting those parameters that are too far from what they have been in the past. However, since MPLE maximizes likelihood of both the data and the priors, we would expect too high of an $m$ to ignore actual data at the cost of the fit. Our strategy is to determine a level of prior information, $\Delta m$, that will correct loss of information due to a change in data, $\Delta y$. In effect, we aim to find $m$ such that NLP stays on an isocline of likelihood under reasonable perturbations in $y_j$. To measure the sensitivity of NLP to changes in the data and the priors, we take a derivative of NLP with respect to $y_j$ and prior weighting:

$$d\text{NLP} = \frac{\partial(\text{NLP})}{\partial y_j} \Delta y_j + \frac{\partial(\text{NLP})}{\partial m} \Delta m. \quad (2.6)$$

For no information loss $\partial\text{NLP} = 0$, and it follows that

$$0 = \frac{\partial(\text{NLP})}{\partial y_j} \Delta y_j + \frac{\partial(\text{NLP})}{\partial m} \Delta m,$$
where we can approximate $\Delta y$, but $\Delta m$ is unknown. Rearranging:

$$\frac{\partial \text{NLP}}{\partial m} \Delta m = -\frac{\partial \text{NLP}}{\partial y_j} \Delta y_j.$$ 

Let $A = \frac{\partial \text{NLP}}{\partial y_j} \Delta y_j$ and $B = \frac{\partial \text{NLP}}{\partial m}$. We can then approximate:

$$\Delta m \approx \left| \frac{\partial \text{NLP}}{\partial y_j} \Delta y_j \right| \approx \frac{A}{B}.$$  (2.7)

First, we derive an estimated change in NLP with respect to a change in data.

$$A \approx \sum_{j=1}^{n} \Delta y_j \frac{\partial \text{NLP}}{\partial y_j} = \sum_{j=1}^{n} \Delta y_j \cdot \frac{\partial}{\partial y_j} \left[ \left( \sqrt{y_j} - \sqrt{Y_j} \right)^2 / w_j * 2\sigma^2 + \frac{1}{2} \log(2\pi\sigma^2) \right].$$

that is

$$\sum_{j=1}^{n} \Delta y_j \frac{\partial \text{NLP}}{\partial y_j} = \frac{1}{2\sigma^2} \sum_{j=1}^{n} \left[ \Delta y_j \cdot \frac{\sqrt{Y_j} - \sqrt{Y_j}}{w_j / \sqrt{y_j}} \right].$$  (2.8)

Recall that backfit is $\sqrt{y_j} / \sqrt{Y_j}$ for root-transformed data. If the model predicts the data with a backfit between 0.80 and 1.20, we have a best case model-data interaction. Then

$$0.8 \lesssim \frac{\sqrt{Y_j}}{\sqrt{y_j}} \lesssim 1.20$$

Also, recall that $w_j = (\sqrt{y_j} - \sqrt{y_j})^2$, normalized by $c$. Assuming $w_j \propto \Delta y_j$, the normalization gives

$$\sum_{j=1}^{n} \Delta y_j / w_j \approx 1.$$  (2.10)

This reduces Equation 2.8 to

$$\Delta y_j \frac{\partial \text{NLP}}{\partial y_j} = \frac{1}{2\sigma^2} \sum_{j=1}^{n} 1 \cdot \left[ 1 - \frac{\sqrt{Y_j}}{\sqrt{y_j}} \right].$$  (2.9)

With $n = 16$ and $\sqrt{y_j} / \sqrt{Y_j} \approx 0.80$, we simplify Equation 2.9 to

$$A \approx \Delta y_j \frac{\partial \text{NLP}}{\partial y_j} \approx \frac{16}{2\sigma^2} \cdot [1 - 0.80] \approx \frac{0.20 \cdot 16}{2\sigma^2}.$$  (2.11)
This gives the sensitivity of NLP to data with normally acceptable backfit values. Next, we estimate the change in NLP with respect to a change in prior information, that is

\[ \Delta m \ast B \approx \Delta m \sum_{k_p=1}^{5} \left[ \frac{(\theta_{k_p}^* - \hat{\theta}_{k_p})^2}{2\nu_{k_p}^2} + \frac{1}{2} \log(2\pi\nu_{k_p}^2) \right] . \] (2.11)

We assume that the estimated parameters, \( \hat{\theta}_{k_p} \), are within one standard deviation of the mean, \( \theta_{k_p}^* \), for data with a reasonable backfit so that \( |\theta_{k_p}^* - \hat{\theta}_{k_p}| \lesssim \nu_{k_p} \) so

\[ B \approx \left[ \frac{5}{2} + \frac{1}{2} \sum_{k_p=1}^{5} \log(2\pi\nu_{k_p}^2) \right] . \]

Now we are able to make a priori estimates of \( \Delta m \). Using Analyte 1 as a test, we take the values of \( \nu_{k_p} \) and \( \theta_{k_p}^* \) from prior estimates seen in Appendix B and roughly estimate \( B \approx 3.40 \). Then, with an averaged \( \sigma \approx 0.26 \), we estimate for Analyte 1

\[ A \approx \Delta y_j \frac{\partial(NLP)}{\partial y_j} \approx \frac{0.20 \ast 16}{2 \ast 0.07} = 24.2. \]

Thus to properly weight the priors according to Equation 2.7

\[ \Delta m \approx \left| \frac{A}{B} \right| \approx \frac{24.2}{3.40} \approx 7. \]

So, in the case of Analyte 1, an \( m \) of 7 should suffice to correct information loss for data near an acceptable backfit. Depending on the severity of information loss due to bad data, we can adjust \( m \) to recover information using prior knowledge. The goal in weighting prior information is to increase \( m \) until quality of fit measures improve, using the prior probabilities of parameters to guide parameter selection. We first calculate a \( \Delta m \) for each analyte individually following the above example for Analyte 1. Then, in a range of \( \Delta m \) as \( 0 \leq m \leq 2 \ast \Delta m \) we look for the \( m \) that gives the maximum percentage of acceptable backfits, for each replicate series and each analyte. This range of \( m \), when \( m \) causes improvement, for each individual set of replicate series in an analyte, \( m_{\text{min}} \) and \( m_{\text{max}} \), will be compared with \( \Delta m \) estimates for each analyte to determine whether \( \Delta m \) accurately predicts the search region for the appropriate weight of prior information. As a validation measure we calculate the fraction of replicate series with and without \( m \) that achieve a full range of of quantification, indicated by an LLOQ of \( x_8 \).
2.4 Developing a Standard Design Using Fewer Sample Wells

Cost to the consumer depends directly on the dilution scheme as wells allocated for the standard become unavailable for sample testing. To reduce the number of wells used by the standard, we can determine if all the data points in a seven part serial dilution are necessary. Rather than redesigning a new dilution scheme, we decide to keep a seven part serial dilution and only load a subset of these dilutions into the wells for analysis. If parametrization using prior probabilities guides the parameter selection to the same values using less points, we can essentially remove those points from the dilution series and in an effort to save on cost.

Using a series of replicate serial dilutions on Analyte 1, we fit a nominal curve using all data points with a NLL minimization routine. Then, we remove both replicates of one concentration value, $x_j$, and refit the curve without these observations. Aiming to determine which removals allow for each $Y_j$ to remain close to the original $y_j$, we measure the CD of all points, including those that were not used for fitting. The deleted replicate points are replaced and the process repeated until each of the replicate points have been removed once. Using the maximum CD as a measure of fit, we find the combination of removals that consistently produces a CD nearest to the maximum CD. As it is unlikely that removing a pair of points at the same $x_j$ will consistently lead to a global maximum CD, we look for the top range of CD measurements and find the single $x_j$ for which the pair of replicate removals consistently appear in the top range across all thirteen replicate series. Lastly, we extend this process onto all replicate series of each of the six analytes. The single pair of removed replicate points that appear in every top-range CD then becomes the dilution we choose to leave unloaded.

We continue this process to include double, triple, quadruple, and even quintuple replicate point removals, iterating through all possible removal combinations and storing the CD for each iteration with the goal to determine the best combinations of points to remove. We classify the best combination of points to remove as the combination of points that produces a CD in the highest range of CDs for all replicate series of all analytes. Once the design is determined, we fit the new dilution series using both MLE and MPLE. In using MPLE, we select an $m$ in the range of $0 \leq m \leq 2 \times \Delta m$ for each replicate series of each analyte that maximizes all backfit values, including the removed data that do not influence parameterization. When $m$ improves the acceptable backfit
percentage, the range that gives the best percentage, $m_{\text{min}}$ to $m_{\text{max}}$ is compared with $\Delta m$ to test the accuracy of the search region predicted by $\Delta m$. We test the quality of fit on the designs for both methods using a percentage of points, including those that no longer influence the fit, in an acceptable backfit. We also test the quality of fit in the low concentrations by the percentage of replicates per analyte with a full range of quantification.
Chapter 3
Results

3.1 Outlier Detection

With every analyte in our unaltered data set, we use the outlier detection algorithm to locate any data points that qualify as outliers at $\alpha = 0.01$, $\alpha = 0.05$, and $\alpha = 0.10$. For each set of replicates we measure the percentage of data in the 0.80 to 1.20 backfit range before and after outlier detection and removal. When $\alpha = 0.01$ no data points were identified as outliers in any of the replicates. As a result, we see no change in the percentage of acceptable fit points as shown in Figure 3.1 below. If we select a higher $\alpha$ the algorithm becomes more exclusive. When $\alpha = 0.05$ a minimal amount of data are excluded from a few of the replicates which generally leads to an increase in acceptable fitted points as shown in Figure 3.2. When $\alpha$ is too high, the algorithm becomes highly exclusive. For example, when we set $\alpha = 0.10$, several of the replicates

![Figure 3.1](image.png)

Figure 3.1: Percentage of data points with a backfit between 0.80 and 1.20 both before and after outlier detection and removal. We see no change in these percentages as no points were removed at this level.
Figure 3.2: Percentage of data points with a backfit between 0.80 and 1.20 both before and after outlier detection and removal. Several replicates contained one to two detected outliers, which were removed.

Figure 3.3: Percentage of data points with a backfit between 0.80 and 1.20 both before and after outlier detection and removal. The percentages improve due to the loss of several data points in each replicate.

lose between two and four data points each. The improvement in acceptably fitted points, shown in Figure 3.3, are due to the volume of loss in poorly fitted data points, not an improvement of fit. We notice a slight decline in the acceptable percentage for Analyte 4 which happens as a result of unnecessary removals. To illustrate the effect of unnecessary removals, we show the percentage of acceptable backfits for all sixteen data points in series eight of this analyte across increasing $\alpha$ in Figure 3.4. Initially, all the data were in an acceptable backfit as shown. When the algorithm removes the
Figure 3.4: Percentage of data points with a backfit between 0.80 and 1.20 in series eight of Analyte 4. The initial percentage is already at a maximum when $\alpha = 0$ and when $\alpha = 0.05$ the percentage degrades on remaining data points.

The first point at $\alpha = 0.05$ the “outlying” $y_j$ fell out of an acceptable backfit as it no longer influences the parameterization of this series. As $\alpha$ increases, $y_j$ that had an acceptable backfit to start are removed and the percentage of acceptable backfit continues to fall. Here, we can clearly see that removing outliers on a “good” data set is not useful.

3.2 Weighting Prior Information

Numerically, the best results for selecting an $m$ align with the isocline approximations of $m$ seen in Chapter 2. The $\Delta m$ appropriate for each analyte is shown in Table 3.1.

<table>
<thead>
<tr>
<th>Ana. 1</th>
<th>Ana. 2</th>
<th>Ana. 3</th>
<th>Ana. 4</th>
<th>Ana. 5</th>
<th>Ana. 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$</td>
<td>0.3</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>$A$</td>
<td>24.2</td>
<td>64.0</td>
<td>36.1</td>
<td>70.0</td>
<td>21.3</td>
</tr>
<tr>
<td>$B$</td>
<td>2.4</td>
<td>1.0</td>
<td>7.8</td>
<td>3.0</td>
<td>2.8</td>
</tr>
<tr>
<td>$\Delta m$</td>
<td>10</td>
<td>64</td>
<td>5</td>
<td>23</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 3.1: Approximations on $\Delta m$ by analyte according to a nullcline analysis from Section 2.3. These approximations are averaged over thirteen replicates for each analyte.
We used a range of $0 \leq m \leq 2 \times \Delta m$ specific to each analyte to locate an optimum $m$. For each replicate series of a standard for each analyte, an $m$ is chosen giving a set

$$\{m : m \in [m_{\text{min}}, m_{\text{max}}]\}.$$ 

where the lowest $m$ that improved the acceptable backfit percentage for a replicate series is denoted as $m_{\text{min}}$ while the maximum $m$ that improved backfit percentages in a replicate series is denoted as $m_{\text{max}}$. In any given replicate series for an analyte, the $m$ used is either between $m_{\text{min}}$ and $m_{\text{max}}$ or $m = 0$, as not all replicate series showed backfit percentage improvement using $m$. When $m = 0$, no prior information is incorporated into the posterior likelihood. We expect MPLE to achieve the same results as MLE when $m = 0$, but if some $m$ improves the fit of the model then we would expect MPLE to have better results than MLE.

<table>
<thead>
<tr>
<th>Ana. 1</th>
<th>Ana. 2</th>
<th>Ana. 3</th>
<th>Ana. 4</th>
<th>Ana. 5</th>
<th>Ana. 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal MLE</td>
<td>0.9471</td>
<td>0.9519</td>
<td>0.9952</td>
<td>0.9808</td>
<td>0.9904</td>
</tr>
<tr>
<td>Nominal MPLE</td>
<td>0.9567</td>
<td>0.9663</td>
<td>0.9952</td>
<td>0.9952</td>
<td>0.9904</td>
</tr>
<tr>
<td>$m_{\text{min}}$</td>
<td>5</td>
<td>1</td>
<td>none</td>
<td>0.5</td>
<td>none</td>
</tr>
<tr>
<td>$m_{\text{max}}$</td>
<td>8</td>
<td>6.5</td>
<td>none</td>
<td>1</td>
<td>none</td>
</tr>
<tr>
<td>$\Delta m$</td>
<td>7</td>
<td>40</td>
<td>5</td>
<td>23</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 3.2: Percentage of data in an acceptable backfit for the nominal MLE and the nominal MPLE curve fitting procedures. Also shown are the minimum and maximum $m$ that improved acceptable backfit percentages in some replicate series by analyte and our nullcline estimate of $\Delta m$.

The estimate of $\Delta m$ uses the nullcline analysis from Section 2.3 and in effect determines the likely $m_{\text{max}}$. We see in Table 3.2 that the range of $m$ do not exceed the maximum of our test range, $2 \times \Delta m$. These results indicate that a nullcline analysis to determine $m_{\text{max}}$, that is $\Delta m$, will provide a range of possible $m$ values, especially when applying a buffer of $2 \times \Delta m$. As a validation measure, the percentage of replicates per analyte with a full range of of quantification, indicated by an LLOQ of the lowest $x_j$, shows slight improvement, as seen in Table 3.3.
Table 3.3: The percentage of replicate series with an LLOQ at the lowest concentration. Percentages are out of thirteen replicate series for each unperturbed analyte.

<table>
<thead>
<tr>
<th></th>
<th>Ana. 1</th>
<th>Ana. 2</th>
<th>Ana. 3</th>
<th>Ana. 4</th>
<th>Ana. 5</th>
<th>Ana. 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal MLE</td>
<td>0.69</td>
<td>0.77</td>
<td>0.92</td>
<td>0.92</td>
<td>0.92</td>
<td>0.69</td>
</tr>
<tr>
<td>Nominal MPLE</td>
<td>0.77</td>
<td>0.85</td>
<td>0.92</td>
<td>0.92</td>
<td>0.92</td>
<td>0.92</td>
</tr>
</tbody>
</table>

3.3 Standard Design with Fewer Sample Wells

In removing one, two, three, four, and five replicate pairs in every possible combination, we took the best dilution scheme as a seven part dilution less the pairs that were present in a range of maximum CD value across all replicate series and all analytes. To illustrate the magnitude of difference in fit while removing replicates, we use the density plot shown in Figure 3.5, averaged across thirteen replicate series of Analyte 1.

The colors represent the distance from the wSSE determined curve to the data as a SSE. High SSEs mean the data are far from the curve. This plot is constructed from a symmetric matrix where the first pair of removals lie along the abscissa and the second pair of removals lie along the ordinate. Since we cannot remove more than one pair of each replicate, the diagonal of this matrix is null. Visually demonstrating the removal of all possible combinations of two replicate pairs, we can infer that removing $x_1$ and $x_2$ at the same time cause the most detriment to the fit, with $\text{SSE} = 53,708.39$. We can also
determine that the best set of replicate pairs to remove are $x_6$ and $x_8$, with SSE = 76.89, closely followed by the next best set of removals, $x_6$ and $x_7$, with SSE = 79.59. The difference between the best set of removals and the worst set of removals is extreme and apparent.

We use a normalized metric, CD, to locate the best set of replicate removals. For the removal of three replicate pairs, we found that the highest, consistent CD occurs with the removal of replicate concentrations $x_4$, $x_6$, and $x_7$. In four replicate pair removal we selected to remove replicate concentrations $x_2$, $x_4$, $x_6$, and $x_8$. Finally, in five replicate pair removal we selected to remove replicate concentrations $x_2$, $x_4$, $x_5$, $x_7$, and $x_8$. To relate the CD to backfits, we removed these replicate points from all replicates of all analytes and measured the nominal backfits and post-removal backfits of all sixteen points. The results for three, four, and five replicate removals, summarized in Table 3.5, show that weighting prior information, MPLE, tends to increase the percentage of acceptable backfits in these sparse design cases. The maximum and minimum $m$ used to attain these results, by analyte, are shown in Table 3.6. We compare these results against the nominal $\Delta m$ to estimate the precision of $\Delta m$ in estimating an appropriate range on $m$. Again we have included $m_{\text{max}}$ in the range of $m$ tested between $0 \leq m \leq 2 \ast \Delta m$ for all

<table>
<thead>
<tr>
<th>Dilution Number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 Removals:</td>
<td>XX</td>
<td>XX</td>
<td>XX</td>
<td>OO</td>
<td>XX</td>
<td>OO</td>
<td>OO</td>
<td>XX</td>
</tr>
<tr>
<td>4 Removals:</td>
<td>XX</td>
<td>OO</td>
<td>XX</td>
<td>OO</td>
<td>XX</td>
<td>OO</td>
<td>XX</td>
<td>OO</td>
</tr>
<tr>
<td>5 Removals:</td>
<td>XX</td>
<td>OO</td>
<td>XX</td>
<td>OO</td>
<td>XX</td>
<td>OO</td>
<td>OO</td>
<td>OO</td>
</tr>
</tbody>
</table>

Table 3.4: New standard designs with three, four, and five replicate point removals. The removed replicates are indicated by 'OO'. The replicates that remain and parameterize the model are indicated by 'XX'.

<table>
<thead>
<tr>
<th>Ana.</th>
<th>Ana. 2</th>
<th>Ana. 3</th>
<th>Ana. 4</th>
<th>Ana. 5</th>
<th>Ana. 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal MLE</td>
<td>0.9471</td>
<td>0.9519</td>
<td>0.9952</td>
<td>0.9808</td>
<td>0.9904</td>
</tr>
<tr>
<td>10 Point MPLE</td>
<td>0.9375</td>
<td>0.9567</td>
<td>0.9808</td>
<td>0.9904</td>
<td>0.9808</td>
</tr>
<tr>
<td>10 Point MLE</td>
<td>0.9135</td>
<td>0.8990</td>
<td>0.9375</td>
<td>0.9183</td>
<td>0.9423</td>
</tr>
<tr>
<td>8 Point MPLE</td>
<td>0.9231</td>
<td>0.9663</td>
<td>0.9712</td>
<td>0.9904</td>
<td>0.9904</td>
</tr>
<tr>
<td>8 Point MLE</td>
<td>0.9135</td>
<td>0.9231</td>
<td>0.9615</td>
<td>0.9423</td>
<td>0.9567</td>
</tr>
<tr>
<td>6 Point MPLE</td>
<td>0.8846</td>
<td>0.9231</td>
<td>0.9711</td>
<td>0.9904</td>
<td>0.9663</td>
</tr>
<tr>
<td>6 Point MLE</td>
<td>0.8750</td>
<td>0.9134</td>
<td>0.9182</td>
<td>0.9567</td>
<td>0.9375</td>
</tr>
</tbody>
</table>

Table 3.5: Percentage of data in an acceptable backfit for the nominal design with all seven dilutions and each of the reduced dilution schemes. Percentages are out of sixteen data points averaged across thirteen replicates for each of the six analytes.
new designs with the exception of the six point design. The extreme \( m \) used in the six point design indicate that the high volume of data removed is not appropriate for these analytes with our methods. The eight point design, however, still uses \( m \) within the specified range and we also begin to see \( m_{\text{max}} \) closer to the approximated \( \Delta m \) across analytes. In both the eight and ten point designs, our method of \( m \) estimation appears reasonable.

Each replicate series of each analyte use different optimal \( m \). Figure 3.6 illustrates the different parameterized curves, with and without \( m \), for the same series of Analyte 6 replicates across the three new designs. The six point design uses \( m_{\text{optimum}} = 24.4 \), the eight point design uses \( m_{\text{optimum}} = 0 \) and the ten point design uses \( m_{\text{optimum}} = 0.2 \). To measure the effect of point removal on the low concentrations, we measure the LLOQ for each replicate series. The percentage of each set of replicates with an LLOQ at the lowest initially measured concentration is summarized by analyte in Table 3.7. More data is usually better in both MLE and MPLE, but to outweigh the cost of each data point we found that it is possible to use less data in setting the standard curve. We ended the removal process with a total loss of ten data points, leaving six data points.

### Table 3.6: The minimum and maximum \( m \) used in testing by analyte. A replicate series may have an optimal fit with \( m = 0 \) or an \( m \) between the minimum and maximum shown.

<table>
<thead>
<tr>
<th></th>
<th>Ana. 1</th>
<th>Ana. 2</th>
<th>Ana. 3</th>
<th>Ana. 4</th>
<th>Ana. 5</th>
<th>Ana. 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 Point ( m_{\text{min}} )</td>
<td>1.4</td>
<td>0.2</td>
<td>0.2</td>
<td>0.4</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>10 Point ( m_{\text{max}} )</td>
<td>6.4</td>
<td>0.4</td>
<td>2.2</td>
<td>7.6</td>
<td>1.6</td>
<td>9.2</td>
</tr>
<tr>
<td>8 Point ( m_{\text{min}} )</td>
<td>4.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.6</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>8 Point ( m_{\text{max}} )</td>
<td>7.8</td>
<td>3.4</td>
<td>2.2</td>
<td>36.6</td>
<td>1.6</td>
<td>10.6</td>
</tr>
<tr>
<td>6 Point ( m_{\text{min}} )</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>6 Point ( m_{\text{max}} )</td>
<td>86</td>
<td>37.2</td>
<td>84.8</td>
<td>0.2</td>
<td>0.2</td>
<td>44.2</td>
</tr>
<tr>
<td>Nominal ( \Delta m )</td>
<td>7</td>
<td>40</td>
<td>5</td>
<td>23</td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>

### Table 3.7: Percentage of replicates with an LLOQ at the lowest initially measured concentration summarized by analyte. Percentages are shown for both MLE and MPLE across each of three, four, and five pair replicate point removals.

<table>
<thead>
<tr>
<th></th>
<th>Ana. 1</th>
<th>Ana. 2</th>
<th>Ana. 3</th>
<th>Ana. 4</th>
<th>Ana. 5</th>
<th>Ana. 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal MLE</td>
<td>0.69</td>
<td>0.77</td>
<td>0.92</td>
<td>0.92</td>
<td>0.92</td>
<td>0.69</td>
</tr>
<tr>
<td>10 Point MLE</td>
<td>0.54</td>
<td>0.77</td>
<td>0.85</td>
<td>0.92</td>
<td>0.85</td>
<td>0.70</td>
</tr>
<tr>
<td>10 Point MLE</td>
<td>0.31</td>
<td>0.38</td>
<td>0.54</td>
<td>0.46</td>
<td>0.46</td>
<td>0.38</td>
</tr>
<tr>
<td>8 Point MLE</td>
<td>0.54</td>
<td>0.77</td>
<td>0.76</td>
<td>0.92</td>
<td>0.85</td>
<td>0.69</td>
</tr>
<tr>
<td>8 Point MLE</td>
<td>0.31</td>
<td>0.54</td>
<td>0.46</td>
<td>0.69</td>
<td>0.69</td>
<td>0.62</td>
</tr>
<tr>
<td>6 Point MLE</td>
<td>0.62</td>
<td>0.31</td>
<td>0.62</td>
<td>0.85</td>
<td>0.85</td>
<td>0.92</td>
</tr>
<tr>
<td>6 Point MLE</td>
<td>0.15</td>
<td>0.46</td>
<td>0.46</td>
<td>0.69</td>
<td>0.62</td>
<td>0.77</td>
</tr>
</tbody>
</table>
Figure 3.6: Parameterized curves for each of the three designs using fewer sample wells for the standard where the removed $x_j$ are shown as 'O' and the remaining $x_j$ used for parameterization are shown as 'X'. Each curve is fit with $m = 0$ (black) and $m = m_{\text{optimum}}$ (green). The difference of fit in the ten point design is nearly undetectable, with a small change in the fit near $x_8$ increasing the LLOQ. The eight point design does not benefit from $m$ in this series. The six point design shows a noticeable difference in fit when using $m$, where $m = m_{\text{optimum}}$ again improves the LLOQ.
for parameterization purposes. Since we need at least five observations to parameterize a five parameter curve, we are disinclined to remove more than five pairs of replicate points. In principle, any five data points would give five parameters, but ideally the five points would need to be at five different concentrations. The six point design, using only three concentrations, does not parameterize as well as the ten and eight point designs, even with prior information.
Chapter 4  
Discussion and Conclusion

4.1 Summary

ELISA is a powerful quantitative tool with predictive abilities completely determined by the accuracy of a five-parameter curve fit to an analyte standard. Parameterizing the standard uses 12-16 out of 96 available wells that are then lost to testing samples, inferring cost on the ELISA consumer and limiting the amount of standard data for fitting to two replicates at each concentration. In addition to limited sample wells, fitting the standard curve is highly subject to fitting error resultant from outlying data, which is impossible to detect with only two replicates, and spurious minima surrounding optimal parameters. If the standard curve cannot properly approximate samples of known concentration, then the inversion of the standard will not accurately quantify samples of unknown concentration. Both inaccurate parameterization and lengthy standard dilution series add unnecessary expense to ELISA consumers. Our methods address outlying data, poor parameterization, and standard design efficiency using a Bayesian perspective to introduce prior information and influence the probability of parameters in fitting the five-parameter model.

The prior information we utilize is the data collected on an analyte during quality control checks of plate development. Parameterizing all of the quality control replicates allows us to develop a collection of best-fit parameters and summarize a likely parameter distribution by mean and variance. In effect, the priors cause the curve fitting algorithm to select parameters that are similar to past parameterization instances rather than parameters that are influenced only by the current data set and whatever problems it might have.

Our results clearly indicate that outliers can be effectively identified and removed
by sampling posterior distributions of intensities developed from prior parameter distributions, choosing a credibility level of $\alpha = 0.05$ and removing data that fall outside of the $(\alpha, 1 - \alpha)$ credibility boundary. The outlier detection algorithm successfully identified outliers in all of the analytes tested at differing $\alpha$. Using threshold values of $\alpha = [0.01, 0.05]$, outliers were identified after $\alpha = 0.01$ and were removed more often at $\alpha = 0.10$ as expected.

Next, the prior information is weighted to correct for the negative effects of poor quality data. We add an extra parameter, $m$, to control the weighting of prior parameter probability and penalize for the selection of parameters that are unlikely according to prior information. We successfully define the range of $m$ to test using an isocline analysis. Using a level of prior information in a pre-determined range, $0 \leq m \leq 2 \times \Delta m$, our results show a 2% overall backfit improvement in 80% of the analytes tested and an LLOQ improvement in between 8% and 23% of these cases. We notice that weighting prior information is particularly helpful in rescuing low concentration accuracy.

Lastly, we use prior probabilities to guide parameter selections using fewer standard wells in an effort to save on cost to the consumer. We assume a seven-part serial dilution and skip those dilutions whose removal degrades the curve fit the least across all analytes tested. A ten point standard design constructed by retaining the first through third dilutions, the fifth, and the seventh dilution results in over 89% fidelity when compared with a full eight point standard design. Using prior information in the ten point design improves fidelity to between 93% and 99%. An eight point standard design with the first, third, fifth, and seventh dilutions, showed over 91% fidelity, improving to between 92% and 99% with the use of prior information. A six point design with the first, third, and sixth dilutions, showed over 87% fidelity, improving to between 88% and 99% fidelity with the use of prior information. The LLOQ in the ten point design shows between 31% and 54% of cases with full quantification, improving to 54% to 92% with the use of prior information. The eight point design showed between 31% and 69% of cases in full quantification, improving with prior information to between 54% and 92%. Lastly, the six point design showed between 15% and 77% cases in full quantification, improving to 31% to 92% with the use of prior information. The nullcline analysis held true to determine a range on $m$ in the ten and eight point designs, but the six point design, likely losing too many points, did not select $m$ within the pre-determined range. According to these results, we recommend the eight point standard design.
Using prior information requires relatively little effort. Data produced during quality control efforts provide the prior information and only parameter means and standard deviations are required. In addition to the hyperparameters, the weighting scheme, if used, must be determined \textit{a priori}, so that using less data does not influence precision at low concentrations. It would be beneficial, over time, to develop better or even discrete posterior distributions that may allow for the elimination of weighting and data transformations entirely.

4.2 An Example Application

In developing outlier detection and a weighting on prior information, we have provided an unsupervised pathway for rescuing plates that face serious data discrepancies. For example, if a replicate of the standard misses a wash, the resultant pixel intensities will read lower than that of the properly washed replicate. We have simulated this example on the eight point design seen in Figure 4.1.

\textbf{Figure 4.1:} A simulated missed wash replicate where the first attempt at fitting this standard with wNLL provides two of eight points in an acceptable backfit. Outliers are detected in the first and second concentrations, then the fit improves to acceptable backfits in five of six remaining points. We capture all six data points to an acceptable backfit by tuning the weight on prior information.

First, the data is fit with a wNLL and the backfit values read two of eight points in the necessary 0.80 to 1.20 backfit. Since this acceptable backfit percentage is clearly unacceptable, we detect outliers at $\alpha = 0.05$ with hopes to improve the backfits. Our
algorithm detected outliers at $x_1$ and $x_2$ in the missed wash replicate. After we remove these data from the parameterization routine, we see an improved fit with five of six remaining points in an acceptable backfit. However, the LLOQ = $x_3$ as the fit has not yet achieved full quantification. We tune the parameter $m$ to reach a weighting of prior information that most improves the fit and we then arrive at full quantification with all remaining data in an acceptable backfit. With this combination of methods we avoid losing all of the information on this plate and require no extra input from the user.

4.3 Future Work

Our Bayesian perspective relies entirely upon a small pool of prior information. Our methods would improve by having more samples to develop parameter distributions, also allowing for the ability to use non-normal parameter distributions. If we had a large enough pool of prior samples, we could use a discrete distribution, rather than making a distribution assumption that introduces more error into the model, improving results that incorporate prior information.

A combination of outlier detection and MPLE may prove beneficial in improving measures of fit while removing only a minima of data, especially since the algorithm had a tendency to identify low concentration replicates as outliers. We also note that a new standard design might be different for specific analytes, rather than across all analytes collectively. We suggest applying our design method to each analyte individually, looking for the combination of removals that provides the absolute maximum CD specific to each analyte.

Lastly, developing a priori $m$ values would also assist in characterizing the curves that use very few data points. While we were able to increment $m$ and calculate the backfit values for the removed data in addition to the remaining data, in practice, backfits cannot be calculated for data aside from the data used for parameterization. Then, rather than choosing $m$ with only the available data and few backfits, we would select $m$ from prior information where the estimates on prior weights are better informed.
Bibliography


Appendix A: Abbreviated Data

The data used in this project appears in columns of replicates by rows of serial dilution. The first column of recorded values are the known concentration levels and the following columns are measured pixel intensities for each of six analytes.

<table>
<thead>
<tr>
<th>Analyte 1 (spot 2)</th>
<th>Concentration, pg/ml</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>150</td>
</tr>
<tr>
<td>S2</td>
<td>545</td>
</tr>
<tr>
<td>S3</td>
<td>181.66666666666666</td>
</tr>
<tr>
<td>S4</td>
<td>20 1515151515151515</td>
</tr>
<tr>
<td>S5</td>
<td>0.123456789012345678</td>
</tr>
<tr>
<td>S6</td>
<td>2.223456789012345678</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Analyte 2 (spot 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
</tr>
<tr>
<td>S2</td>
</tr>
<tr>
<td>S3</td>
</tr>
<tr>
<td>S4</td>
</tr>
<tr>
<td>S5</td>
</tr>
<tr>
<td>S6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Analyte 3 (spot 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
</tr>
<tr>
<td>S2</td>
</tr>
<tr>
<td>S3</td>
</tr>
<tr>
<td>S4</td>
</tr>
<tr>
<td>S5</td>
</tr>
<tr>
<td>S6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Analyte 4 (spot 6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
</tr>
<tr>
<td>S2</td>
</tr>
<tr>
<td>S3</td>
</tr>
<tr>
<td>S4</td>
</tr>
<tr>
<td>S5</td>
</tr>
<tr>
<td>S6</td>
</tr>
<tr>
<td>Analyte (spot 7)</td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>41407</td>
</tr>
<tr>
<td>41408</td>
</tr>
<tr>
<td>41409</td>
</tr>
<tr>
<td>41410</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Analyte (spot 8)</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>S6</th>
<th>S7</th>
<th>S8</th>
</tr>
</thead>
<tbody>
<tr>
<td>41407</td>
<td>6366</td>
<td>6220</td>
<td>104</td>
<td>34</td>
<td>11</td>
<td>3.9</td>
<td>41407</td>
<td></td>
</tr>
<tr>
<td>41408</td>
<td>5221</td>
<td>5488</td>
<td>3133</td>
<td>17595</td>
<td>6595</td>
<td>2198</td>
<td>14</td>
<td>41408</td>
</tr>
<tr>
<td>41409</td>
<td>5498</td>
<td>3484</td>
<td>104</td>
<td>13649</td>
<td>6220</td>
<td>58617</td>
<td>14</td>
<td>41409</td>
</tr>
<tr>
<td>41410</td>
<td>5498</td>
<td>3133</td>
<td>104</td>
<td>11784</td>
<td>5220</td>
<td>556</td>
<td>27.5</td>
<td>41410</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Analyte (spot 8)</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>S6</th>
<th>S7</th>
<th>S8</th>
</tr>
</thead>
<tbody>
<tr>
<td>41407</td>
<td>6220</td>
<td>5960</td>
<td>5488</td>
<td>6366</td>
<td>2620</td>
<td>5488</td>
<td>3133</td>
<td>41407</td>
</tr>
<tr>
<td>41408</td>
<td>5498</td>
<td>4685</td>
<td>3133</td>
<td>5960</td>
<td>6366</td>
<td>5221</td>
<td>3484</td>
<td>41408</td>
</tr>
<tr>
<td>41409</td>
<td>5498</td>
<td>4685</td>
<td>3133</td>
<td>13649</td>
<td>6366</td>
<td>6220</td>
<td>5221</td>
<td>41409</td>
</tr>
<tr>
<td>41410</td>
<td>5960</td>
<td>4685</td>
<td>3484</td>
<td>11784</td>
<td>6366</td>
<td>5221</td>
<td>3484</td>
<td>41410</td>
</tr>
</tbody>
</table>
Appendix B: Prior Estimates Used

Prior estimates for each of the six analytes tested. These estimates are from parameterizing each of the analyte replicate series and keeping on the best-fit instances. From the pool of best fits, we use the mean and variance, specific to each analyte.

<table>
<thead>
<tr>
<th></th>
<th>Ana. 1</th>
<th>Ana. 2</th>
<th>Ana. 3</th>
<th>Ana. 4</th>
<th>Ana. 5</th>
<th>Ana. 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a^*$</td>
<td>252.726</td>
<td>253.199</td>
<td>267.187</td>
<td>255.864</td>
<td>255.504</td>
<td>261.047</td>
</tr>
<tr>
<td>$\nu_a$</td>
<td>3.1987</td>
<td>2.1152</td>
<td>19.8783</td>
<td>2.8269</td>
<td>2.6207</td>
<td>3.2657</td>
</tr>
<tr>
<td>$b^*$</td>
<td>-13.2596</td>
<td>-11.634</td>
<td>-10.014</td>
<td>-10.902</td>
<td>-11.243</td>
<td>-10.104</td>
</tr>
<tr>
<td>$\nu_b$</td>
<td>1.5801</td>
<td>0.8659</td>
<td>2.2970</td>
<td>2.2163</td>
<td>1.5039</td>
<td>1.3980</td>
</tr>
<tr>
<td>$c^*$</td>
<td>2.7439</td>
<td>2.8547</td>
<td>4.9263</td>
<td>3.0063</td>
<td>3.2237</td>
<td>3.6381</td>
</tr>
<tr>
<td>$\nu_c$</td>
<td>0.0284</td>
<td>0.0384</td>
<td>0.0873</td>
<td>0.0302</td>
<td>0.0645</td>
<td>0.0770</td>
</tr>
<tr>
<td>$d^*$</td>
<td>7.8035</td>
<td>7.7545</td>
<td>5.1356</td>
<td>10.1315</td>
<td>8.9692</td>
<td>7.5337</td>
</tr>
<tr>
<td>$\nu_d$</td>
<td>1.4255</td>
<td>1.3885</td>
<td>2.5716</td>
<td>1.9272</td>
<td>1.4011</td>
<td>1.5440</td>
</tr>
<tr>
<td>$g^*$</td>
<td>0.2114</td>
<td>0.1712</td>
<td>0.3743</td>
<td>0.1580</td>
<td>0.2056</td>
<td>0.2727</td>
</tr>
<tr>
<td>$\nu_g$</td>
<td>0.0441</td>
<td>0.0237</td>
<td>0.2021</td>
<td>0.0463</td>
<td>0.0402</td>
<td>0.0445</td>
</tr>
</tbody>
</table>
Appendix C: MATLAB Code

C.1: Parameterization Function

This function calculates an error associated with fitting the five parameter curve to an input of standard analyte data. The function is used in conjunction with MATLAB’s `fminsearch` to minimize the error by changing the parameter values. The analyte input is the column of known concentrations and and two replicate columns of observed pixel intensities. A vector of prior means initializes the parameter values, the tuning parameter is set by the user, and a vector of a priori weights averaged across the analyte are input into the function. Details on the input format are listed in the code comments below.

```matlab
function [err] = wNPL( analyte, parameters, priors, reps, m, weights )

% % % % % % % % % % % % % % % % % % % % % %
% calculates the error associated with fitting a logistic curve to analyte data.
% % Inputs:
% analyte: analyte of interest in columns [ concentration ; replicate_1:replicate_n ]
% parameters: vector of starting guesses for five parameter logistic parameters
% priors: column vector of priors
% reps: column numbers of analyte to fit
% m: numeric prior tuning parameter
% weights: column vector of apriori weights of the same length as the analyte replicate
% % Output:
% err: Error to minimize
```

d = parameters(1); % d, the estimated response at infinite concentration
a = parameters(2); % a, the estimated response at zero concentration
b = parameters(3); % b, the slope factor
c = parameters(4); % c, the mid-range concentration
g = parameters(5); % g, asymmetry factor

% Name priors
sigma_d = priors(1);
sigma_a = priors(2);
sigma_b = priors(3);
sigma_c = priors(4);
sigma_g = priors(5);
dstar = priors(6);
astar = priors(7);
bstar = priors(8);
cstar = priors(9);
gstar = priors(10);

% Make size adjustments
% dilution becomes n x n
dilution = log(analyte(:,1))*ones(1,2);
% make one column vector from all columns
dilution = dilution(:);
% set analyte replicates
an1 = analyte(:,reps);
% make one column vector from both replicate columns
pixint = an1(:);

% Calculate weighting constant
constant = 1/(sum(1./weights));

% Predicted pixel intensity values for current parameters and dilutions
pixintpred = d + (a - d)./((1 + abs((dilution./c).^b)).^g);
% Calculate sigma
[nobs] = length(pixint);
sigma = sqrt(1/nobs*sum(constant.*( (pixintpred-pixint).^2 ./weights ));

%Calculate error
err = nobs/2*log(2*pi*sigma^2)...
    +nobs*sum(constant.*( (pixintpred-pixint).^2 ./weights )/(2*sigma^2)...)...
    + M*(0.5*log(2*pi*sigma_d^2)+(d-dstar)^2/(2*sigma_d^2)+... + 0.5*log(2*pi*sigma_a^2)+(a-astar)^2/(2*sigma_a^2)+... + 0.5*log(2*pi*sigma_c^2)+(c-cstar)^2/(2*sigma_c^2)+... + 0.5*log(2*pi*sigma_g^2)+(g-gstar)^2/(2*sigma_g^2)+... + 0.5*log(2*pi*sigma_b^2)+(b-bstar)^2/(2*sigma_b^2));

end

C.2: Outlier Detection Algorithm

The outlier detection code detects outliers in a data set at a given threshold. The inputs into the function require two replicates of the analyte, previously determined “best fit” parameters, the number of samples to simulate, the analyte known concentrations, and the desired cutoff threshold. The user must first fit the data before identifying any outliers. A small function to generate curves from the randomly sampled parameters is called as five_parameter_plotter in the code below. Also refer to the code for details on input and output formats.

function [R1Outlier, R2Outlier, WhichOne] = Detect_Outliers_Posteriorly(reps, sse parms, sample_size, analyte, threshold)

%Detect_Outliers_Posteriorly is a function that detects outliers in an analyte dataset
%using a family of fits derived from prior information

%Inputs:
%reps: data columns to detect outliers on formatted as a:b for rep a rep b
%sse_parms: pool of prior parameters
%sample_size: number of family of fits, 10000 is a good place to start
%analyte: data
%threshold: a cutoff "alpha" for determining outliers between 0.03 and 0.07
%Outputs:
%R1Outlier: Outliers identified in replicate one
%R2Outlier: Outliers identified in replicate two
%WhichOne: Concentration number of identified outliers

% Name priors
sigma_d = priors(1);
sigma_a = priors(2);
sigma_b = priors(3);
sigma_c = priors(4);
sigma_g = priors(5);
dstar = priors(6);
astar = priors(7);
bstar = priors(8);
cstar = priors(9);
gstar = priors(10);

% Develop credible distribution
d = abs(normrnd(dstar, sigma_d, [sample_size,1]));
a = abs(normrnd(astar, sigma_a, [sample_size,1]));
b = normrnd(bstar, sigma_b, [sample_size,1]));
c = abs(normrnd(cstar, sigma_c, [sample_size,1]));
g = abs(normrnd(gstar, sigma_g, [sample_size,1]));
intensities = zeros(length(d),length(analyte(:,1)));
for z=1:length(d)
    intensities(z,:) = five_parameter_plotter(analyte,[d(z), a(z), b(z), c(z), g(z)]);
end

% Initialize value-range and outlier storage vectors
LVal = zeros(length(analyte(:,1),1);
UVal = zeros(length(analyte(:,1),1);
R1Outlier = zeros(length(analyte(:,1),1);
R2Outlier = zeros(length(analyte(:,1),1);

%Develop credible distribution exclusion rules
for p = 1:length(analyte(:,1)
    [N,x] = hist(intensities(:,p), sample_size);
    CN = cumsum(N)/sum(N);
    %Lower Value Cutoff
    LVR = find(CN >= threshold/2 & CN <= (threshold+threshold/2));
    [aL] = linear_regression(x(LVR),CN(LVR));
    %Upper Value Cutoff
    UVR = find(CN >= (1-(threshold+threshold/2)) & CN <=
               ((1-threshold/2)));
    [aU] = linear_regression(x(UVR),CN(UVR));
    LVal(p) = (threshold-aL(2))/aL(1);
    UVal(p) = ((1-threshold)-aU(2))/aU(1);
    %Find values outside cutoff
    if ((analyte(p,reps(1)) < LVal(p)) || (analyte(p,reps(1)) >
        UVal(p)))
        R1Outlier(p) = analyte(p,reps(1));
    end
    if ((analyte(p,reps(2)) < LVal(p)) || (analyte(p,reps(2)) >
        UVal(p)))
        R2Outlier(p) = analyte(p,reps(2));
    end
end

%Exclude outliers and store in a vector for output
I1 = find(R1Outlier);
I2 = find(R2Outlier)+8;
WhichOne = [I1; I2];