2014

Dynamic Bayesian Approaches to the Statistical Calibration Problem

Derick Lorenzo Rivers
Virginia Commonwealth University

Dynamic Bayesian Approaches to the Statistical Calibration Problem

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

by
Derick L. Rivers
Doctor of Philosophy

Director: Dr. Edward L. Boone, Associate Professor
Department of Statistical Sciences and Operations Research

Virginia Commonwealth University
Richmond, Virginia
December 2014

Dedication

To the greatest grandmother on earth, Ruby Rivers.....

Acknowledgments

I would like to thank my advisor Edward L. Boone, my committee members David J. Edwards, QiQi Lu, James E. Mays and Nitai Mukhopadhyay whose generous help, encouragement, and expert guidance were responsible for the completion of this dissertation. I would also like to thank D'Arcy Mays and Jason Merrick for giving me the opportunity to follow my dreams and providing continuous support and guidance along the way.
I would like to give a special thanks to Mrs. Ruth Lattimore for introducing me to the language of Statistics and instilling in me a desire to learn.
I am truly grateful to my family and friends who had to put up with me during the last few years.

Contents

List of Tables

List of Figures

Abstract

DYNAMIC BAYESIAN APPROACHES TO THE STATISTICAL CALIBRATION PROBLEM

By Derick L. Rivers, Ph.D.

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2014.

Director: Dr. Edward L. Boone, Associate Professor, Department of Statistical Sciences and Operations Research

The problem of statistical calibration of a measuring instrument can be framed both in a statistical context as well as in an engineering context. In the first, the problem is dealt with by distinguishing between the "classical" approach and the "inverse" regression approach. Both of these models are static models and are used to estimate "exact" measurements from measurements that are affected by error. In the engineering context, the variables of interest are considered to be taken at the time at which you observe it. The Bayesian time series analysis method of Dynamic Linear Models (DLM) can be used to monitor the evolution of the measures, thus introducing a *dynamic* approach to statistical calibration. The research presented employs the use of Bayesian methodology to perform statistical calibration. The DLM framework is used to capture the time-varying parameters that may be changing or drifting over time. Dynamic based approaches to the linear, nonlinear, and multivariate calibration problem are presented in this dissertation. Simulation studies are conducted where the dynamic models are compared to some well known "static" calibration approaches in the literature from both the frequentist and Bayesian perspectives. Applications to microwave radiometry are given.

Chapter 1

Introduction

Statistical calibration is a type of *reverse* prediction problem that is related to common regression problems. This sort of problem is sometimes called *inverse regression* due to its resemblance of linear regression. With linear regression, it is assumed that for a random vector $\mathbf{Y}$ and a given value of another variable $\mathbf{X} = \mathbf{x}$ which is not random, the conditional expectation is linear in $\mathbf{x}$:

$$E[\mathbf{Y}|\mathbf{X} = \mathbf{x}] = \mathbf{X}\boldsymbol{\beta} \tag{1.1}$$

where $\mathbf{Y}$ is an $(n \times 1)$ vector of responses, $\mathbf{X}$ is an $(n \times (k+1))$ model matrix, $\boldsymbol{\beta}$ is a $((k+1) \times 1)$ vector of unknown fixed parameters. The i^{th} row of the model matrix $\mathbf{X}$ will be denoted as $\mathbf{x}_i'$.

In regression, measurement pairs $(\mathbf{x}_i', y_i)$ are collected. For each measurement pair it is required that linearity of conditional expectation hold for the calculated values, therefore some assumptions are required: one must assume that the $\mathbf{x}_i$ are nonrandom and observed with negligible measurement error, while the ϵ_i, in the error vector $\boldsymbol{\epsilon}$, are uncorrelated normally distributed random variables with mean equal 0 and a constant variance σ^2. Adding the aforementioned assumptions to equation (1.1) we get the random observation vector $\mathbf{Y}$, given by

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{1.2}$$

where $\boldsymbol{\epsilon}$ is a $(n \times 1)$ vector of uncorrelated error terms.

In regression, one foci of interest is on the estimation of the vector of unknown parameters; $\boldsymbol{\beta}$, and possibly on the prediction of future $\hat{\mathbf{Y}}_{i|new}$ values corresponding to fixed $\mathbf{X} = x_{i|new}'$ values. This type of prediction problem is relatively straightforward, due to the fact that a future $\mathbf{Y}_i$ value is a random variable, and the probability statement, $P(\mathbf{Y} = y_i|\mathbf{X} = \mathbf{x}_i')$, about the future value can be made directly.

To illustrate the idea of linear regression, lets consider a simple linear regression (i.e. a single x_i) example provided by Moore et al. (2010) where a researcher measured the heights (in centimeters) of 161 children in Kalama, a village in Egypt. The heights were averaged and recorded each month, with the study lasting several years. The scatter plot of the data

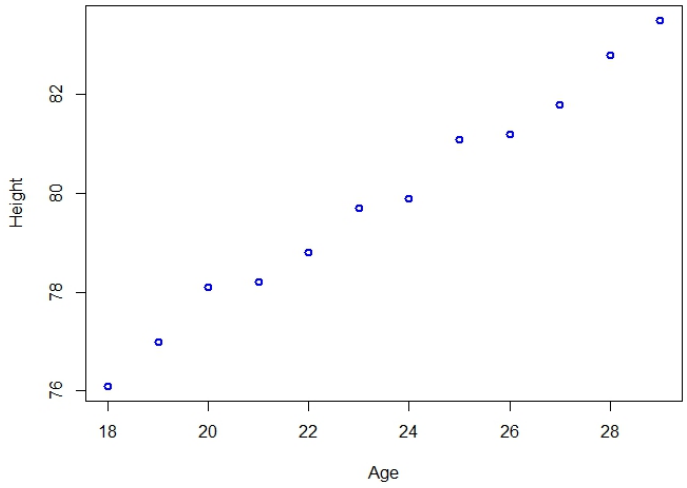


Figure 1.1: *Children Heights vs. Age*

is presented in Figure 1.1 where *age* (in months) is the regressor variable and the observed *average height* is the response variable.

Note that the data in Figure 1.1 is approximately linear. We see as the age increases, the average height increases at an approximately constant rate. The aim of regression is to plot the *Line of Best Fit*, often the *Least Squares Regression Line* by estimating the unknown parameters (β's) in the model.

The *Least Squares Regression Line* is the line that has the minimal sum of the errors squared (SSE), which is the squared difference between the true observed value y_i and the predicted value $\hat{y}_i$ from a given set of data. The sum of the errors squared is defined by Moore et al. (2010) as:

$$SSE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}(y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i))^2. \tag{1.3}$$

Moore et al. (2010) obtain the estimates of the parameters; β_0 and β_1, by taking the derivative of SSE with respect to β_0 and β_1 and setting them to zero giving the following set of equations:

$$\frac{\partial SSE}{\hat{\beta}_0} = 2n\hat{\beta}_0 + 2\hat{\beta}_1 \sum_{i=1}^{n} x_i - 2\sum_{i=1}^{n} y_i = 0 \tag{1.4}$$

$$\frac{\partial SSE}{\hat{\beta}_1} = 2\hat{\beta}_1 \sum_{i=1}^{n} x_i^2 + 2\hat{\beta}_0 \sum_{i=1}^{n} x_i - 2\sum_{i=1}^{n} x_i y_i = 0 \tag{1.5}$$

Solving the normal equations gives the following least square estimates of β_0 and β_1 as:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}, \tag{1.6}$$

and

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})} \tag{1.7}$$

where $\bar{x} = (1/n)\sum_{i=1}^{n} x_i$ and $\bar{y} = (1/n)\sum_{i=1}^{n} y_i$.

For the data points in Figure 1.1, $n = 12$, and it is found from equations (1.6) and (1.7) that $\hat{\beta}_0 = 64.928$ and $\hat{\beta}_1 = 0.635$. Hence, the least squares regression line is $\hat{y} = 64.928 + 0.635x$. This line is given in Figure 1.2.

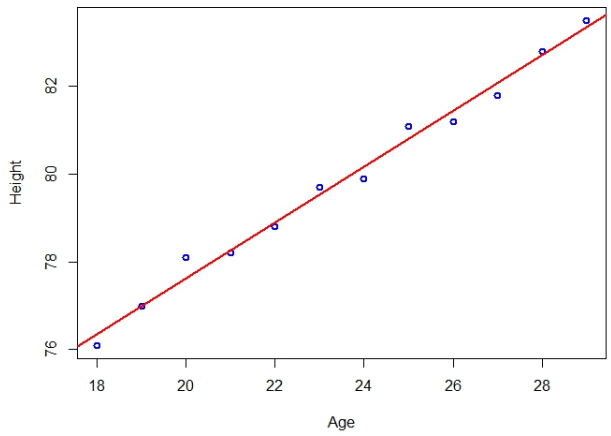


Figure 1.2: *Heights vs. Age with regression line.*

The problem of statistical calibration differs from that of standard regression, in that a primary point of interest is using a known *observed* value of the response, say y_0 to estimate an *unknown* value of the regressor, say x_0. A calibration experiment is carried out in two stages. First a sample of pairs (x_i, y_i), $i = 1, 2, \ldots, n$, of data is observed. At this stage the calibration curve/line is established. Secondly, m observations of the response are observed, all corresponding to a single unknown value of x_0. Since the regressor is not assumed to be a random variable, inferences in the calibration problem are inherently different from inferences in a regression (or prediction) problem.

When using the term calibration, one must distinguish between *comparative* calibration and *absolute* calibration. The term *comparative* calibration is used when one instrument or measurement technique is calibrated against another where neither one is a standard measurement, implying that there is no standard measurement x. Madansky (1959) considers

the comparative calibration problem where the standards and test methods are respectively denoted by ξ and η. It is assumed that $\eta = f(\xi) = \beta_0 + \beta_1\xi$, where β_0 and β_1 are the intercept and slope parameters respectively. In comparative calibration, n pairs of observations (x_i, y_i) are obtained in the first stage where x_i and y_i are the observed values for ξ_i and η_i respectively,

$$y_i = \eta_i + \epsilon_i \qquad i = 1, 2, \ldots, n, \tag{1.8}$$

$$x_i = \xi_i + \delta_i \qquad i = 1, 2, \ldots, n, \tag{1.9}$$

where ϵ_i and δ_i are the error terms. The term *absolute* calibration is considered when $\delta_i = 0$ for all i.

Absolute calibration is defined to be a technique where the measurements, x, are known or taken with negligible error. Williams (1969a) and Lwin and Spiegelman (1986) refer to absolute calibration when imprecise measurements y are related to values of working standards, x. By making the assumption that $\delta_i = 0$ we have

$$y_i = \eta_i + \epsilon_i = f(x_i) + \epsilon_i \qquad i = 1, 2, \ldots, n, \tag{1.10}$$

thus becoming the linear calibration problem

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \qquad i = 1, 2, \ldots, n. \tag{1.11}$$

Aitchison and Dunsmore (1975) and Brown (1982) states that a univariate calibration problem can be classified as being either a *controlled* calibration problem or a *random* (or *natural*) calibration problem. In a controlled calibration problem, it is assumed that the values of the regressor from the first stage of experimentation are fixed by the researcher. On the other hand, in a random calibration problem, the values $x_1, x_2, \ldots, x_n$ of the regressor variable are randomly generated. In the latter case, although the regressor and response are both random variables, they are still treated asymmetrically. It is assumed that the values of the regressor at the first stage of experimentation are precisely measurable, while the measurements of the response variable are subject to error. This asymmetry is defined by the model and error structure which the relationship between the regressor and the response is assumed to have. In this work, we use the term calibration to refer to absolute and controlled calibration; comparative calibration and random calibration are not considered. A review of the literature on statistical calibration is given by Osborne (1991).

.This dissertation is organized as follows. In the remainder of this chapter we consider a literature review of the linear calibration problem from frequentist and Bayesian perspectives, Dynamic Linear Models, and simulation techniques. Chapter 2 contains a novel dynamic approach to the linear statistical calibration problem. An application to microwave radiometry is also considered. Chapter 3 is concerned with extending the dynamic approach in Chapter 2 to the case where there is nonlinearity in the covariate. We present results that can be used to guide researchers in the placement of reference measurements. In Chapter 4, the use of a multi-univariate method is explored. In this case, we model a multi-device calibration scenario by assuming a multivariate response variable and an explanatory variable that is univariate. Finally, in Chapter 5, closing remarks and future work for the methods presented here will be considered.

1.1 The Frequentist Approaches to Calibration

1.1.1 The "Classical" Approach

The first solution to the absolute calibration problem was proposed by Eisenhart (1939) which is now called the *classical* estimator to the calibration problem. Eisenhart assumed that the relationship between x and y was of a simple linear form:

$$E(Y|X = x) = \beta_0 + \beta_1 x.$$

The estimated calibration line of Y on X is given by

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X, \tag{1.12}$$

where $\hat{\beta}_0$ is the estimate of the intercept β_0 and $\hat{\beta}_1$ is the estimate of the slope β_1. The significant characteristic of the work by Eisenhart (1939) is that the line relating Y to X is fit by minimizing the residual sum of squares in the direction of the response Y, using the data collected at the first stage of experimentation. That is, the estimates of β_0 and β_1 are those that minimize

$$SSE_y = \sum_{i=1}^{n} (y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i))^2. \tag{1.13}$$

After the parameter estimates; $\hat{\beta}_0$ and $\hat{\beta}_1$ are found, Eisenhart (1939) inverts equation (1.12) to estimate the unknown regressor value x_0 by using the observed response value y_0:

$$\hat{x}_{0,c} = \frac{y_0 - \hat{\beta}_0}{\hat{\beta}_1} \tag{1.14}$$

where $\hat{\beta}_0$ and $\hat{\beta}_1$ are the least square estimates found by equations (1.6) and (1.7).

This estimator of $x_{0,c}$ is known as the "classical" estimator to the calibration problem. Although this solution does not require a specific distributional assumption about the deviations of the y values from their means, this estimator would be the maximum likelihood estimator of x_0 if independent, identically distributed normal errors were assumed on the y_i's.

Ideally, when a researcher is studying a calibration problem, more than just a point estimate is required. The need for an interval estimate may be required in this case. This "classical" approach to interval estimation has been the subject of several papers because if the slope $\hat{\beta}_1$ is not significantly different from zero the interval is either the whole line or even two disjoint *semi-infinite* lines. A *semi-infinite* line is defined to be bounded in one direction, and unbounded in another. In this case, interval estimation is of no use. To address this problem Berkson (1969) and Shukla (1972) derive asymptotic expressions for the mean squared error (M.S.E) and the bias of $\hat{x}_{0,c}$ given that $|\hat{\beta}_1| > 0$. Brown (1993) describes an

interval estimate that is useful when we are convinced that $\hat{\beta}_1$ is significantly different from zero. Brown's (1993) interval estimate is

$$\frac{y_0 - \hat{\beta}_0}{\hat{\beta}_1}\left(1 + \frac{\hat{\sigma}^2 t^2}{\hat{\beta}_1^2 S_{xx}}\right) \pm \frac{\hat{\sigma} t}{\hat{\beta}_1}\left(1 + \frac{1}{2n} + \frac{(y_0 - \hat{\beta}_0)^2 + \hat{\sigma}^2 t^2}{2\hat{\beta}_1^2 S_{xx}}\right), \tag{1.15}$$

which is similar to Fieller's (1954) solution to the problem of inference about the ratio of two means.
This estimate is an approximate $100(1-\gamma)\%$ interval estimate for x_0, where γ is a small non-negative number close to 0. The estimate of the standard deviation is defined as:

$$\hat{\sigma} = \sqrt{\frac{SSE_y}{n-2}},$$

the sum of the squared differences between each regressor value and the mean of the regressors is:

$$S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2,$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and t is the $100(1-\frac{\gamma}{2})$ percentile of the Student's t distribution with $n-2$ degrees of freedom.

Consider the following example to illustrate the classical approach to statistical calibration. Let the x_i's be known radiometer voltage measurements and let the y_i's represent the corresponding temperature readings. If the voltage measurement of an object with unknown voltage x_0 is desired, we regress Y on to X and use Eq. (1.14) where $\hat{\beta}_0$ and $\hat{\beta}_1$ are found by equations (1.6) and (1.7).

In this example, the values for $\hat{\beta}_0$ and $\hat{\beta}_1$ are -17.58 and 3.93, respectively. If we say the $y_0 = 40$, equation (1.14) yields the estimate:

$$\hat{x}_{0,c} = 14.64$$

with the 95% interval estimate being $(13.46, 15.76)$. The 95% interval estimate was found from equation (1.15).

1.1.2 The "Inverse" Approach

The *inverse* approach to the calibration problem was first presented by Krutchkoff (1967). The model in the *inverse* approach is written as

$$X_i = \phi + \delta Y_i + \epsilon_i^{'} \tag{1.16}$$

where ϕ and δ are the parameters in the linear relationship and the $\epsilon_i^{'}$ are the measurement errors. This model suggest that the regression be done by relating X to Y and minimizing

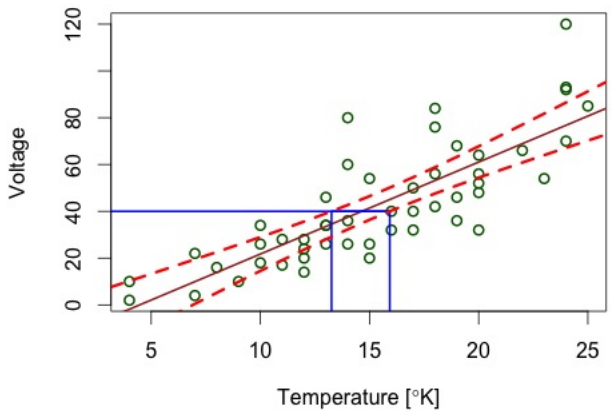


Figure 1.3: *Classical Statistical Calibration Example*

the residual sum of squares in the X direction:

$$SSE_x = \sum_{i=1}^{n} (x_i - (\hat{\phi} + \hat{\delta} y_i))^2. \tag{1.17}$$

Clearly, the slope and intercept parameters are different from those previously discussed. The least square estimates for the intercept term ϕ, and the slope δ, are:

$$\hat{\phi} = \bar{x} - \hat{\delta}\bar{y} \tag{1.18}$$

and

$$\hat{\delta} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(y_i - \bar{y})}, \tag{1.19}$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$.

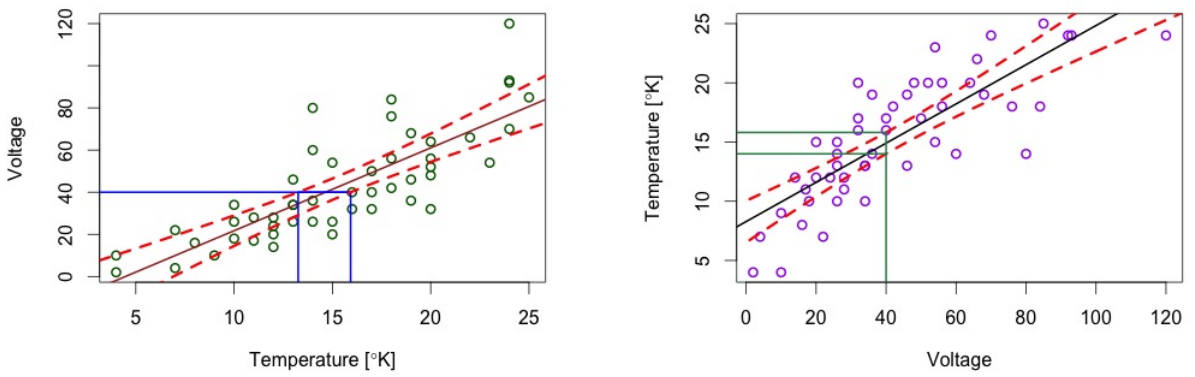
At the second stage of experimentation, x_0 can be estimated directly by substituting y_0 into the fitted equation. We let $\hat{x}_{0,I}$ denote the inverse estimator of x_0:

$$\hat{x}_{0,I} = \hat{\phi} + \hat{\delta} y_0 \tag{1.20}$$

Krutchkoff (1967) based his conclusion on the results of a simulation study, in which he found that the mean squared error of estimation for x_0 was uniformly less for this estimator than for the classical estimator.

Let's use the same example previously illustrated to compare the "inverse" statistical calibration approach to the "classical" approach. Again, we let the x_i's be known radiometer voltage measurements and let the y_i's represent the corresponding temperature readings. If

the voltage measurement of an object with unknown voltage x_0 is desired, this time we regress X on to Y and use equation (2.4) where $\hat{\phi}$ and $\hat{\delta}$ are found by equations (1.18) and (1.19), respectively. The values for $\hat{\phi}$ and $\hat{\delta}$ are found to be 8.28 and 0.17, respectively. If



(a) *"Classical" Approach* **(b)** *"Inverse" Approach*

Figure 1.4: *Scatterplots of "Classical" and "Inverse" Approaches*

we say the $y_0 = 40$, equation (1.20) yields:

$$\hat{x}_{0,I} = 14.91$$

with the 95% interval estimate being $(14.00, 15.81)$. Clearly, in Figure 1.4 we observe that the "classical" approach and the "inverse" approach yield similar results. Even though the results appear to be similar the publication of Krutchkoff's (1967) paper motivated several responses.

Some of these responses criticized Krutchkoff's (1967) conclusion, while others defended it. Williams (1969b) noted that the "inverse" estimator chosen by Krutchkoff (1967) has been derived by least squares on the false assumption that the errors ϵ' are independent of the values of y, an assumption which is in conflict with the original relations postulated. Williams (1969b) agreed that the inverse estimator has a finite variance, however concluded that this was its only merit. He also demonstrated that any unbiased estimator of x_0 must have infinite mean squared error. Based on these facts, he concluded that mean squared error (MSE) is not a suitable criterion for the comparison of the estimators.

Berkson (1969) pointed out that the inverse procedure does not yield consistent estimators for β_0, β_1, and x_0, while the classical procedure does. Berkson (1969) states that it is impossible for an estimate based on an inconsistent statistic to have in all conditions a smaller mean squared error for all values of the estimated quantity. It is pointed out that the smaller mean squared error of the inverse estimate in estimating X from a single observation of Y, noted in the sampling experiment, holds only for a limited range of x. It was noted

by Berkson (1969) that all of the experiments performed by Krutchkoff (1967) had small N, where N was 6 for most of the experiments with $N = 20$ being the largest. He shows for large N, $N \to \infty$, that the estimates, $\hat{\phi}$ and $\hat{\delta}$ are biased. This argument is shown by first understanding for the "classical" estimates, as $N \to \infty$,

$$\hat{\beta}_0 \rightarrow \beta_0, \tag{1.21}$$

$$\hat{\beta}_1 \rightarrow \beta_1, \tag{1.22}$$

meaning that the estimates $\hat{\beta}_0$ and $\hat{\beta}_0$ are consistent estimates. Berkson (1969) goes on to show that as $N \to \infty$, the "inverse" estimates are

$$\tilde{\phi} = -\frac{\hat{\phi}}{\hat{\delta}} \rightarrow \beta_0 - \frac{\sigma^2}{\beta_1 \sigma_x^2}\bar{x} = \beta_0 - f\bar{x}, \tag{1.23}$$

$$\tilde{\delta} = \frac{1}{\hat{\delta}} \rightarrow \beta_1 + \frac{\sigma^2}{\beta_1 \sigma_x^2} = \beta_1 + f, \tag{1.24}$$

where $f = \sigma^2/\beta_1\sigma_x^2$, σ is the standard error of the $\epsilon^{'}$, the error term, and σ_x^2 is the variance of the x_i's used in determining the calibration line. Thus, the estimates $\hat{\phi}$ and $\hat{\delta}$ are not consistent. Finally, he demonstrated via simulation that for some parameter values the classical procedure can be shown to outperform the inverse procedure.

Shukla (1972) stated that when the number of observations used in the calibration experiment is small, then for suitable choice of design of independent variables the "inverse" estimator will yield a smaller mean square error. He followed by stating that in practice when a large number of observations are used for calibration with small error and the unknown x_0 is estimated by large number of observations on Y, it is unlikely that the "inverse" method will be advantageous over the "classical" method except in very trivial cases. In many cases it is not possible to take more than one observation on unknown x_0, and the "inverse" method of estimation is preferable when x_0 lies close to the mean of the design points so far as MSE is taken as the comparison criterion. However, for general purposes it is advisable to prefer an estimator that is consistent for large sample sizes, which suggests the use of the "classical" estimator in the absence of any prior information about the unknown x_0 (Shukla, 1972).

Several other papers followed Krutchkoff (1967). A subsequent study by Krutchkoff (1969) compared the "classical" and "inverse" calibration approaches and demonstrated that for a sufficiently large number of observations, the "classical" method produces lower mean squared error outside the range of calibration. The earlier results by Krutchkoff (1967) remained unchanged for X values within the prescribed calibration range (i.e. $0 \leq X \leq 1$). Martinelle (1970) and Halperin (1970) both supported Berkson (1969) by concluding that the "inverse" estimator of Krutchkoff (1967) would only hold in restrictive cases.

One paper that provided some support to the inverse procedure was written by Lwin and Maritz (1982). Their approach was based on a method referred to as *linear compound*

estimation. This approach seeks constants k_0 and k_1 so as to minimize the quantity

$$\sum_{i=1}^{n} E\left[\zeta(Y_i) - x_i\right]^2, \tag{1.25}$$

where $\zeta(Y_i) = k_0 + k_1 Y$. Lwin and Maritz (1982) showed by using their *linear compound* estimation approach that the classical approach is favored if unbiased estimation of x_0 is required, while the inverse approach is favored if unbiasedness is not required.

1.2 The Bayesian Approaches to Calibration

A discussion of the Bayesian approach to statistical inference is given before proceeding with the Bayesian solutions to the calibration problem. Bayesian statistics is founded on the fundamental Bayes' theorem:

$$\pi(\boldsymbol{\theta}|\mathbf{y}) = \frac{\pi(\boldsymbol{\theta}) f(\mathbf{y}|\boldsymbol{\theta})}{m(\mathbf{y})}$$

where

$$m(\mathbf{y}) = \int_{\Theta} \pi(\boldsymbol{\theta}) f(\mathbf{y}|\boldsymbol{\theta}) d\boldsymbol{\theta}$$

denotes the marginal distribution of $\mathbf{y}$. The Bayes' theorem combines the available summarized *prior* information about the parameter $\boldsymbol{\theta}$ in a *prior distribution*, $\pi(\boldsymbol{\theta})$, with the information provided by the data $\mathbf{y}$ (referred to as the *likelihood function*), $f(\mathbf{y}|\boldsymbol{\theta})$, to produce the required posterior distribution $\pi(\boldsymbol{\theta}|\mathbf{y})$. Computation of the posterior distribution is often easily derived by noting that Bayes' theorem may be simply expressed as

$$\pi(\boldsymbol{\theta}|\mathbf{y}) \propto \pi(\boldsymbol{\theta}) f(\mathbf{y}|\boldsymbol{\theta}),$$

where $\propto$ means "proportional to".

1.2.1 Hoadley's Bayesian Approach to Calibration

The first noted Bayesian solution to the calibration problem was presented by Hoadley (1970) and is considered to be a milestone along the development of statistical calibration theory. His work was motivated by the unanswered question in the Frequentist community of whether β_1 is zero (or close to zero). Hoadley (1970) first considered the usual F-statistic to test the null hypothesis that $\beta_1 = 0$. Hoadley (1970) showed the F-statistic as $F = \hat{\beta}_1^2 S_{xx}/\hat{\sigma}^2$, where

$$\hat{\sigma}^2 = \left\{\sum_{i=1}^{n}\left(y_{1i} - (\hat{\beta}_0 + \hat{\beta}_1 x_i)\right)^2 + \sum_{j=1}^{m}\left(y_{2j} - \bar{y}_2\right)^2\right\}/(n+m-3).$$

where n is the number of (x_i, y_i) calibration pairs from the first phase of the calibration experiment and m is the number of observation measurements taken at y_0.

If $F > F_{\alpha,1,n-2}$, then we reject the null hypothesis which implies the "classical" estimator, $x_{0,C}$ (Eisenhart, 1939), is satisfactorily accurate. However if $F < F_{\alpha,1,n-2}$, we fail to reject the null hypothesis, thus the "classical" estimator is inaccurate. This gave Hoadley (1970) justification for the use of the "inverse" estimator, $x_{0,I}$ (Krutchkoff, 1967).

Hoadley (1970) expressed the *inverse regression* problem formally as follows:

$$y_{1i} = \beta_0 + \beta_1 x_i + \epsilon_{1i}, \qquad i = 1, 2, \ldots, n \tag{1.26}$$

$$y_{2j} = \beta_0 + \beta_1 x_0 + \epsilon_{2j}, \qquad j = 1, 2, \ldots, m, \tag{1.27}$$

where the error terms, ϵ_{1i} and ϵ_{2j}, are mutually independent and identically distributed as $N(0, \sigma^2)$. Hoadley (1970) assumes that the x_i's are known constants and that β_0, β_1, σ^2, and x_0 are unknown. The observations, $y_{11}, y_{12}, \ldots, y_{1n}, y_{21}, y_{22}, \ldots, y_{2m}$ are used to make inferences about x_0. The x_i's are chosen so that

$$\sum_{i=1}^{n} x_i = 0, \quad \text{and} \quad \frac{1}{n}\sum_{i=1}^{n} x_i^2 = 1.$$

Hoadley (1970) use an example to explain the form that the observations take in equation (1.26). Let the x_i's be known weights and the y_{1i}'s represent the corresponding readings off the scale being calibrated. If someone wanted to weigh an object with unknown weight x_0, she may take m different readings $y_{21}, \ldots, y_{2m}$ from the scale, and then use the y_{1i}'s, y_{2j}, and the x_i's to estimate x_0.

The assumption made by Hoadley (1970) was that x_0 is random and *a priori* independent of $\pi(\beta_0, \beta_1, \sigma^2)$, so that the joint prior distribution of $\pi(\beta_0, \beta_1, \sigma^2, x_0) \propto \pi(\beta_0, \beta_1, \sigma^2)\pi(x_0)$. Hoadley (1970) first assumed that $(\beta_0, \beta_1, \sigma^2)$ had a uniform distribution therefore "*noninformative*" prior σ^{-2} was specified as the prior distribution for $\pi(\beta_0, \beta_1, \sigma^2)$, but the prior distribution for x_0 is left unspecified. Bernardo and Smith (1994) defines *noninformative* prior as a prior distribution that would have a minimal effect, relative to the data, on the posterior inference. Eno (1999) simplifies this definition by stating that a *noninformative* prior is a function which is used in place of a subjective (*informative*) prior distribution when little or no prior information is available.

It was shown that for $m = 1$ (one observation at the prediction stage), that if x_0 has a prior density from a Student t distribution with $n - 3$ degrees of freedom, a mean of 0, and a scale parameter

$$\frac{n+1}{n-3},$$

the posterior is

$$\pi(x_0|\mathbf{Data}) \sim t_{n-2}\left(\hat{x}_{0,I}, \left[\frac{n + 1 + (\hat{x}_{0,I})^2/R}{F + n - 2}\right]\right), \tag{1.28}$$

where $R = \frac{F}{F+n-2}$ and $F = \hat{\beta}_1^2 S_{xx}/\hat{\sigma}^2$.

Although Hoadley (1970) studied this prior distribution, he made no claims about the

desirability of resulting posterior inferences, except to note that this might be a reasonable prior distribution to use if there was reason to believe that the unknown x_0 would likely be close to the mean $\bar{x}$ of the regressor values from the first stage of experimentation.

Aitchison and Dunsmore (1975) contributed to the Bayesian perspective by taking a predictive distribution approach to the statistical calibration problem. The calibration distribution (i.e. posterior distribution for x_0) was derived by assuming the ϵ_i are *i.i.d.* $N(0, \sigma^2)$ with a non-informative prior for $(\beta_0 + \beta_1 x_0, \sigma^2)$. Their calibration density function agrees with that of Hoadley (1970) when $m = 1$. For $m > 1$, Aitchison and Dunsmore (1975) conclude that if a tractable prior distribution is chosen for x_0, the posterior density is a non-central Student distribution function with mean $\hat{x}_{0,B}$. The Bayes estimator $\hat{x}_{0,B}$ is defined as:

$$\hat{x}_{0,B} = \frac{\bar{x} + S_{xy}(\bar{y}_0 - \bar{y})}{S_{yy} + \sum_{j=1}^{m}(y_{0,j} - \bar{y}_0)^2}, \tag{1.29}$$

where $S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$ and $S_{yy} = \sum_{i=1}^{n}(y_i - \bar{y})^2$. They argue that the estimator $\hat{x}_{0,B}$ takes in account the variation in the m future observations in such a way that the more variation there is in the observations, the closer the estimate is to the mean of the prior distribution. For more detail when $m > 1$, see Aitchison and Dunsmore (1975).

1.2.2 Hunter and Lamboy Bayesian Approach to Calibration

Hunter and Lamboy (1981) also considered the calibration problem from a Bayesian point of view. The approach of Hunter and Lamboy (1981) is considered to be similar to that of Hoadley (1970) because also assumed the prior to be

$$\pi(\beta_0, \beta_1, \sigma^2, \eta) \propto \sigma^{-2} \tag{1.30}$$

where $\eta = E[y_0|x_0] = \beta_0 + \beta_1 x_0$. The primary difference between their approach and the approach of Hoadley (1970) is that *a priori* they assume that η and $(\beta_0, \beta_1, \sigma^2)$ were independent while Hoadley (1970) assumed *a priori* that x_0 and $(\beta_0, \beta_1, \sigma^2)$ were independent. Hunter and Lamboy's (1981) work was followed by six papers (Hill, 1981; Lwin, 1981; Rosenblatt and Spiegelman, 1981; Easterling, 1981; Orban, 1981; Lawless, 1981). Hill (1981) was probably the most critical of Hunter and Lamboy's (1981) paper. He raised concern about the choice of priors for η and $(\beta_0, \beta_1, \sigma^2)$. Hill (1981) criticized their assumption of independence and the lack of information on the location, shape and behavior of the posterior distribution of x_0. Later, Brown (1982) criticized Hunter and Lamboy's (1981) prior on η by stating that there was no natural generalization of this prior to several future y_0 observations corresponding to different unknown x_0's. See Osborne (1991) for a summary of these papers.

Smith and Corbett (1987) applied Bayesian and maximum likelihood methods to the estimation of the length of a marathon course. The background of their work was the 1984 Summer Olympic games. Smith and Corbett (1987) developed a "static" approach to calibration under the Bayesian and maximum likelihood paradigms. They also make use of a

"dynamic" model to cope with changes in the calibration constants, β_1. By assuming the intercept, β_0, was equal to zero in equation (1), the slope $\beta_{1,t}$ was allowed to change over time, where

$$\beta_{1,t} = \beta_{1,t-1} + \epsilon_t \qquad t = 1, 2, \ldots, T. \tag{1.31}$$

The disturbances ϵ_t were treated as random variables. This approach follows in the spirit of the state-space modeling approaches to time series problems (see e.g. West et al. 1985; West and Harrison, 1997) and also fits in to the Bayesian strategy framework. Smith and Corbett (1987) conclude that the dynamic model suggests a possible general approach to the recalibration problem.

Recent developments have been made by Eno (1999) and Hickey (2006). Eno (1999) further considers the analysis for the linear statistical calibration problem from the Bayesian perspective. His work was based on the utilization of *noninformative* priors.

Eno (1999) considered three such prior distributions:

1. Jeffrey's Prior (Jeffreys, 1946);
2. Reference Prior Method (Bernardo, 1979); (Berger and Bernardo, 1992); and
3. Probability Matching Priors (Tibshirani, 1989).

His posterior results were analyzed and compared with classical inference procedures. Eno (1999) showed that the noninformative prior Bayesian analysis was a strong competitor, yielding posterior inferences that can, in many cases, be correctly interpreted in a frequentist context. The linear statistical calibration problem was extended to polynomial models and multivariate regression models. For these models, Eno (1999) developed noninformative priors, and derived posterior inferences.

Hickey (2006) explored an exhaustive collection of the primary methods used to solve the statistical calibration problem. Through his examination of the methods, Hickey (2006) designed methods of solving the problem by eliciting priors based on prior beliefs and knowledge. He illustrated this approach through an archaeological problem known as the luminescence problem. In this problem a range of elicited priors were considered; each with a different set of assumptions which led to an intrinsic study of robustness.

1.3 The Nonparametric and Nonlinear Approaches to Calibration

There have been several alternative approaches to statistical calibration that go beyond the general assumptions required for parametric methods. In recent years there have been papers that have taken a non-parametric approach to calibration. The work by Lwin and Maritz (1980), Lwin (1981),and Lwin and Maritz (1982) commonly reference papers when

considering non-parametric approaches to the calibration problem. Lwin and Maritz (1980) consider a random calibration experiment in which only the bivariate random variables (X, Y) can be observed. They derived the distribution of X given Y basing estimation of the marginal distribution of X on the sample distribution function. Lwin and Maritz (1982) compared the classical and inverse estimators $\hat{x}_{0,C}$ and $\hat{x}_{0,I}$; respectively by using a compound estimation approach. With this approach, $\hat{x}_{0,C}$ is a linear compound estimator satisfying the criterion of asymptotic unbiasedness, while $\hat{x}_{0,I}$ is a linear compound estimator without the unbiasedness constraint. Their formulation required no specific distributional assumptions. Their approach showed that $\hat{x}_{0,I}$ was superior to $\hat{x}_{0,C}$ only if the current x_0 value was sampled from the same population as previous X values (i.e., $x_1, x_2, \ldots, x_n$), of the calibration experiment. Two calibration papers were presented by Clark (1979) and Clark (1980) that gave particular reference to radiocarbon dating. Clark (1979) assumed the model

$$Y_{ij} = F(x_i) + e_{ij} \quad i = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, m_i, \quad \sum m_i = N, \tag{1.32}$$

where $\{x_i\}$ are known constants, F is some unknown but smooth function, and e_{ij} are uncorrelated random variables with zero mean but constant variance σ^2. The $\{e_{ij}\}$ and $\{x_i\}$ are assumed to be distinct with $x_1 < x_2 \cdots < x_n$. The estimate $\hat{F}$ was chosen by a cross-validation method defined by a family of functions defined by

$$\hat{F}(x) = \mathbf{u}(x)^T \mathbf{Y} \tag{1.33}$$

where $\mathbf{u}(x) = (u_1(x), u_2(x), \ldots, u_n(x))'$ is a $(nx1)$ vector of known functions, possibly indexed by some index parameter β, $Y_{i\cdot} = \sum_i Y_{ij}/m_i (i = 1, 2, \ldots, n)$, and $\mathbf{Y} = (Y_{1\cdot}, Y_{2\cdot}, \ldots, Y_{n\cdot})'$. Clark (1979) considered the construction of simultaneous calibration intervals under the assumption that the estimator $\hat{F}$ has negligible bias. The paper by Clark (1980) considered the equivalent problem of construction of prediction intervals (i.e., intervals for Y given $X = x$) under the assumption that $\hat{F}$ has a non-negligible bias.

Scheffé (1973) greatly extended this approach to calibration via simultaneous intervals. Lechner et al. (1982) combined Scheffé's (1973) calibration approach with linear splines to produce simultaneous calibration intervals for the liquid volume (v) in large nuclear material processing tanks, given the differential pressure (p). Knafl et al. (1984) also considered pressure-volume calibration of a nuclear tank and adopted Scheffé's procedure (1973), with resulting calibration intervals having two associated probabilities α and δ. Their nonparametric approach is more general than Lechner et al. (1982). Osborne (1990) developed a variety of Bayesian non-parametric approaches to calibration. The nonparametric approach of Gruet (1996) combined kernel and robust estimation techniques. Gruet (1996) illustrated the proposed technique on a radiocarbon dating problem, where first robust point-wise estimates of the parameters of interest were obtained. Second, Gruet (1996) found asymptotic simultaneous tolerance regions for many unknown values of the quantity to be calibrated. For more on nonparametric approaches to calibration see Misquitta and Ruymgaart (2005) and Walker et al. (2010).

1.4 Multivariate Statistical Calibration

Just as simple linear regression can be extended to multiple regression and multivariate regression, the linear calibration problem can be extended to more general settings. The *multivariate linear calibration problem* is similar to the univariate calibration problem, but in the multivariate case it is assumed that q response variables depend linearly on p regressor variables. As before, the data for the calibration experiment are collected in two stages. At the first stage, n observations $(\mathbf{x}_i, \mathbf{y}_i)$ are made. Here, each $\mathbf{x}_i$ is a $p \times 1$ vector, and each $\mathbf{y}_i$ is a $q \times 1$ vector. It is assumed that each element of $\mathbf{x}_i$ is precisely determined, while the elements of $\mathbf{y}_i$ are measurements which are subject to error. At the second stage of data collection, one or more observations of $\mathbf{y}$ are made, and interest centers on the determination of the unknown $\mathbf{x}_0$ vector giving rise to these observations. A new concern comes to light as we extend the univariate calibration problem to the multivariate case, concerning the relative sizes of q and p. Specifically, if the unknown value $\mathbf{x}_0$ is to be uniquely determined, we require that the number of responses q be at least as large as the number of regressors p.

Multivariate calibration was studied extensively by Brown (1982). In his paper, procedures for multivariate calibration that are comparable to the *classical* and *inverse* procedures for univariate calibration were given. Here the term multivariate is intended to mean multivariate in both the X's and the Y's. Suppose there are n observations in the *controlled* calibration experiment, q response variables $Y_1, Y_2, \ldots, Y_q$ and p explanatory variables $X_1, X_2, \ldots, X_p$ and $q \geq p$. Then an model for the first stage of the calibration experiment is given by

$$\mathbf{Y} = \mathbf{1}\boldsymbol{\alpha}' + \mathbf{X}\mathbf{B} + \mathbf{E}, \tag{1.34}$$

where $\mathbf{Y}$ and $\mathbf{E}$ are $(n \times q)$ random matrices, $\mathbf{X}$ is a $n \times p$ matrix of fixed constants and $\mathbf{1}$ is a $(n \times 1)$ vector of ones. $\mathbf{B}$ is a $p \times q$ matrix of unknown parameters, and $\boldsymbol{\alpha}$ is a $q \times 1$ vector of unknown parameters.

Brown (1982) assumes that

$$\sum_{i=1}^{n} x_{ij} = 0, \qquad \frac{1}{n}\sum_{i=1}^{n} x_{ij}^2 = 1, \qquad j = 1, \ldots, p,$$

which means that the columns of $\mathbf{X}$ are standardized and have average sum of squares equal to one. The classical estimator $\hat{\mathbf{B}}$ is given by

$$\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \quad \text{and} \quad \hat{\boldsymbol{\alpha}} = \bar{\mathbf{y}}. \tag{1.35}$$

The model for the prediction experiment is given by

$$\mathbf{y}_0 = \mathbf{1}\boldsymbol{\alpha}' + \mathbf{1}\mathbf{x}_0'\mathbf{B} + \mathbf{E}^{\star}, \tag{1.36}$$

where $\mathbf{y}_0$ and $\mathbf{E}^{\star}$ are $(m \times q)$ random matrices, $\mathbf{x}_0$ is a $p \times 1$ vector of unknown values and $\mathbf{1}$ is a $(m \times 1)$ vector of ones. It is wished to draw inferences about $\mathbf{x}_0$. If $\mathbf{e}_i'$ is the i^{th} row of $\mathbf{E}$, it is assumed that $E(\mathbf{e}_i) = \mathbf{0}$, $E(\mathbf{e}_i\mathbf{e}_i^T) = \boldsymbol{\Gamma}$ and $\mathbf{e}_i \sim N(\mathbf{0}, \boldsymbol{\Gamma})$ for $i = 1, 2, \ldots, n$. If $\mathbf{e}_j^{\star\prime}$ is

the j^{th} row of $\mathbf{E}^{\star}$, the $\mathbf{e}_j^{\star}$ satisfy the above also and it is assumed that they are independent of the $\mathbf{e}_i^{'}$ (Brown 1982; Özyurt and Erar 2003).

It was also shown by using the *classical* approach similar to Eisenhart (1939) that a $100(1-\gamma)\%$ confidence region for $\mathbf{x}_0$ is

$$(\mathbf{y}_0-\hat{\boldsymbol{\alpha}}-\hat{\mathbf{B}}'\mathbf{x}_0)'\mathbf{S}^{-1}(\mathbf{y}_0-\hat{\boldsymbol{\alpha}}-\hat{\mathbf{B}}'\mathbf{x}_0)/_{\sigma^2(\mathbf{x}_0)} \leq \frac{q}{v}F_{\gamma,q,v}, \tag{1.37}$$

where $\mathbf{S}$ is a $(q \times q)$ matrix given by $\mathbf{S} = \hat{\mathbf{E}}'\hat{\mathbf{E}} = (\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}})'(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}})$ with $v = n - p - q$ degrees of freedom and

$$\sigma^2(\mathbf{x}_0) = \frac{1}{m} + \frac{1}{n} + \mathbf{x}_0'(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_0 \tag{1.38}$$

and $F_{\gamma,q,v}$ is the upper $100(1-\gamma)\%$ point of the standard F distribution on q and v degrees of freedom.

Thus the estimation of the unknown vector $\mathbf{x}_0$ is obtained in the form

$$\hat{\mathbf{x}}_0 = (\hat{\mathbf{B}}\mathbf{S}^{-1}\hat{\mathbf{B}}')^{-1}\hat{\mathbf{B}}\mathbf{S}^{-1}(\mathbf{y}_0 - \bar{\mathbf{y}}) \tag{1.39}$$

Brown (1982) also obtains an estimation of $\mathbf{x}_0$ using an inverse estimation method like that of Krutchkoff (1967) where the model is

$$\hat{\mathbf{X}} = \mathbf{Y}\hat{\mathbf{B}}_k \tag{1.40}$$

and $(\mathbf{X}-\hat{\mathbf{X}})'(\mathbf{X}-\hat{\mathbf{X}})$ is minimized using the least squares method. The least squares estimate of $\mathbf{B}_k$ is

$$\mathbf{B}_k = (\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{Y}'\mathbf{X}. \tag{1.41}$$

For a given $1 \times p$ dimensioned observation $\mathbf{y}_0$, $\hat{\mathbf{x}}_{0,k}$ is given by

$$\hat{\mathbf{x}}_{0,k} = \mathbf{X}'\mathbf{Y}(\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{y}_0'. \tag{1.42}$$

In addition, Brown (1982) provided a multivariate calibration extension to the Bayesian approach to the calibration problem by Hoadley (1970). In this Bayesian solution, Brown (1982) used a noninformative prior for all model parameters except the unknown $\mathbf{x}_0$,

$$\pi(\mathbf{B}, \boldsymbol{\alpha}, \boldsymbol{\Gamma}, \mathbf{X}_0) = \pi(\mathbf{B}, \boldsymbol{\alpha}, \boldsymbol{\Gamma})\pi(\mathbf{X}_0), \tag{1.43}$$

where a Jeffrey's invariant prior (Jeffreys, 1946) is assumed,

$$\pi(\mathbf{B}, \boldsymbol{\alpha}, \boldsymbol{\Gamma}) \propto |\boldsymbol{\Gamma}|^{-(q+1)} \tag{1.44}$$

which is the multivariate equivalent to the noninformative prior for $(\alpha, \beta_0, \sigma^2)$ i.e. $\pi(\alpha, \beta_0, \sigma^2) \propto \sigma^{-2}$ proposed by Hoadley (1970). He studied a certain multivariate Student's t distribution as the prior for $\mathbf{x}_0$. As was the case in the earlier work of Hoadley (1970), this prior was studied because the resulting mode of the marginal posterior distribution of $\mathbf{x}_0$ is equal to the inverse estimator of $\mathbf{x}_0$. Brown did not propose any noninformative prior distributions

for $\mathbf{x}_0$.

Several papers followed Brown (1982) that noted that there were a wide variety of approaches to multivariate calibration. More recently Naes et al. (1986) compared the multiple linear regression (MLR), ridge regression (RR), principal component regression (PCR) and partial least squares regression (PLSR) approaches with particular reference to the calibration of near infra-red (NIR) instruments. Brown and Sundberg (1987) considered an approach to multivariate calibration which involves the profile or maximum relative likelihood (Kalbfleisch and Sprott, 1970). An extension to Brown's (1982) multivariate calibration approach is proposed by Fox (1989). His work covers the situation where one is interested in calibrating for an unknown q-vector $\mathbf{X}$ on the basis of an observed p-vector $\mathbf{Y}$ given that $k \geq 1$ components are fixed in advance. Sundberg and Brown (1989) examined the case of there being more variables than observations ($n < p+q+1$) when assuming a standard multivariate linear regression model in equation 4.1. By way of an example, they used the NIR data of Fearn (1983). du Plessis and van der Merwe (1995) developed a Bayesian approach to multivariate and conditional calibration problem. They present a multivariate Bayesian approach akin to that of the univariate calibration approach of Hunter and Lamboy (1981). A Bayesian competitor of the conditional procedure derived by Fox (1989) is also given by du Plessis and van der Merwe (1995). Using the data of Brown (1982), Fox (1989), and du Plessis and van der Merwe (1995), a comparative analysis was done by Özyurt and Erar (2003) that studied how the conditional calibration technique of Fox (1989) and the classical and inverse multivariate techniques of Brown (1982) perform when outliers are present.

1.5 General Dynamic Linear Models

Mathematical and statistical modeling of time series processes is based on classes of *dynamic models*, where the term *dynamic* refers to changes in such processes due to the passage of time (West and Harrison, 1997). The principles employed in Bayesian forecasting and dynamic modeling involve

- parametric models with meaningful dynamic parameters;
- a probabilistic representation of information about parameters;
- a sequential model definition utilizing conditional independence;
- robust conditionally independent model components;
- forecasts derived as probability distributions;
- a facility for incorporating expert information;
- model quality control.

The Dynamic Linear Model (DLM) of West and Harrison (1997) can be represented as a system of equations specifying how observations of a process are stochastically dependent on the current process state and can be represented by how the process parameters evolve in time. The system of equations captures the inherent process dynamics with the stochastic elements modeled by random shocks or disturbances. The DLM is a general class of linear models flexible enough to represent most real time series processes. The model is stated in terms of discrete, equally spaced intervals of time although it is possible to extend it to unequal intervals.

The formulation of the DLM must follow certain principles for Bayesian forecasting and dynamic modeling (Pole et al., 1994). Suppose that at a certain time $t-1$ all the relevant information available up to that time is denoted as D_{t-1}, i.e. D_{t-1} is the set of all observations up to time $t-1$. From the modelers/forecaster point of view interest lies in the forecast value of some scalar quantity, denoted by Y_t with observed values of this quantity as y_t. It follows that $D_t = \{Y_t, D_{t-1}\}$. At time $t-1$ a meaningful parameterization such that all historical information relevant to predicting future observations is contained in a vector denoted $\boldsymbol{\theta}_{t-1}$. This relevant information is represented in terms of a probability distribution $(\boldsymbol{\theta}_{t-1}|D_{t-1})$ such that given D_{t-1}, $(\boldsymbol{\theta}_{t-1}|D_{t-1})$ is sufficient for predicting the future. The parameter vector $(\boldsymbol{\theta}_{t-1})$ must be meaningful, dynamic and changeable to allow incorporation of expert information from the decision makers and from other influential factors outside the system. Current information can then be related to the future via some derived predictive distribution $(Y_{t+k}|D_{t-1}), k = 0, 1, 2, \ldots$. This derivation is via specification of a sequential parametric relation $(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}, D_{t-1})$ together with $(Y_{t+k}|\boldsymbol{\theta}_{t-1}, D_{t-1})$. In combination with $(\boldsymbol{\theta}_{t-1}|D_{t-1})$ these distributions enable derivation of a full joint forecast Student t-distribution. The main property enabling effective dynamic modeling is *conditional independence*, which can be stated generally as follows: given the present state $\boldsymbol{\theta}_t$, the present observation y_t and the future observation y_{t+m} for $m \geq 1$ are independent of the past observation y_{t-1}.

For a time point t, $t = 1, 2, \ldots$, let $M_j, j = 1, 2, \ldots$, represent a model for the underlying time series Y_t with observation y_t. Let $\boldsymbol{\theta}_t$ be a $(r \times 1)$ vector of unknown regression parameters at time t, with p being the number of parameters in the model. The distribution of the underlying time series conditional on the regression parameters and the distribution of the parameters conditional on the previous parameter state can be characterized by a set of quadruples

$$\{\mathbf{F}, \mathbf{G}, V, \mathbf{W}\}_t = \{\mathbf{F_t}, \mathbf{G_t}, V_t, \mathbf{W_t}\} \tag{1.45}$$

for each time t, where $\mathbf{F}_t$ is a known $(n \times r)$ design matrix, $\mathbf{G}_t$ is a known $(r \times r)$ system matrix, V_t is the observation variance and $\mathbf{W}_t$ is the evolution variance matrix.

This quadruple defines the model relating $\mathbf{Y}_t$ to the $(n \times 1)$ parameter vector $\boldsymbol{\theta}_t$ at time t, and the $\boldsymbol{\theta}_t$ sequence through time, via the sequentially specified distributions

$$(Y_t|\boldsymbol{\theta}_t) \sim N[\mathbf{F}_t\boldsymbol{\theta}_t, V_t] \tag{1.46}$$

and

$$(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}) \sim N[\mathbf{G}_t\boldsymbol{\theta}_{t-1}, \mathbf{W}_t]. \tag{1.47}$$

West and Harrison (1997) define for each t, the general univariate DLM by:

$$\begin{aligned} \text{Observation equation :} \quad & Y_t = \mathbf{F}_t\boldsymbol{\theta}_t + \nu_t, \qquad \nu_t \sim N[0, V_t] \\ \text{System equation :} \quad & \boldsymbol{\theta}_t = \mathbf{G}_t\boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \qquad \omega_t \sim N[\mathbf{0}, \mathbf{W}_t] \\ \text{Initial information :} \quad & (\boldsymbol{\theta}_0|D_0) \sim N[\mathbf{m_0}, \mathbf{C_0}], \end{aligned} \tag{1.48}$$

for some prior moments m_0 and C_0. The variances ν_t and $\boldsymbol{\omega}_t$ are independent normally distributed random vectors with zero mean and time varying variance $V_t = E[\nu_t^2]$ and variance-covariance matrices $\mathbf{W}_t = E[\omega_\mathbf{t}\ \omega_\mathbf{t}^{'}]$ respectively at time t.

Sequential updating of the vector of unknown regression parameters, $\boldsymbol{\theta}_{t-1}$, at time $t-1$, to $\boldsymbol{\theta}_t$ at time t occurs by Kalman filtering (Kalman 1960; West and Harrison 1997). This process is carried out by first letting D_t represent all the knowledge from the past up to time t. Inferences about $\boldsymbol{\theta}_t$, $(\boldsymbol{\theta}_t|D_t)$, also known as the *State of Nature*, given all of the knowledge available, D_t, can be carried out in terms of probability densities through the use of Bayes' Theorem. Bayes' Theorem can be expressed as

$$p(\text{State of Nature}|\text{Data}) \propto p(\text{Data}|\text{State of Nature}) \times p(\text{State of Nature}),$$

$$p(\boldsymbol{\theta}_t|D_t) \propto p(Y_t|\boldsymbol{\theta}_t, D_{t-1}) \times p(\boldsymbol{\theta}_t|D_{t-1}). \tag{1.49}$$

Based on the previous stated theorem Harrison and Stevens (1976); West and Harrison (1997) outline the one-step forecast and posterior distributions for each time t as follows:

(a) Posterior at $t-1$: For some mean $\mathbf{m}_{t-1}$ and variance $\mathbf{C}_{t-1}$,
$(\boldsymbol{\theta}_{t-1}|D_{t-1}) \sim N[\mathbf{m}_{t-1}, \mathbf{C}_{t-1}]$.

(b) Prior at time t: $(\boldsymbol{\theta}_t|D_{t-1}) \sim N[\mathbf{a}_t, \mathbf{R}_t]$, where
$\mathbf{a}_t = \mathbf{G}_t\mathbf{m}_{t-1}$ and $\mathbf{R}_t = \mathbf{G}_t\mathbf{C}_{t-1}\mathbf{G}_t^{'} + \mathbf{W}_t$.

(c) One-step forecast: $(Y_t|D_{t-1}) \sim N[f_t, Q_t]$, where
$f_t = \mathbf{F}_t^{'}\mathbf{a}_t$ and $Q_t = \mathbf{F}_t^{'}\mathbf{R}_t\mathbf{F}_t + V_t$.

(d) Posterior at time t: $(\boldsymbol{\theta}_t|D_t) \sim N[\mathbf{m}_t, \mathbf{C}_t]$, with
$\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_t e_t$ and $\mathbf{C}_t = \mathbf{R}_t - \mathbf{A}_t Q_t \mathbf{A}_t^{'}$,
where
$\mathbf{A}_t = \mathbf{R}_t\mathbf{F}_t/Q_t$ and $e_t = Y_t - f_t$.

For the proof by induction using multivariate normal distribution theory refer to West and Harrison (1997).

1.5.1 Dynamic Linear Regression Models

An important case to note is when a constant term is included in the model, such as an intercept term. The result is a straight line regression on $X = X_t$ specified by $\mathbf{F}_t = (1, X_t)'$ and $\boldsymbol{\theta}_t = (\beta_{0t}, \beta_{1t})'$. Then

$$\begin{aligned} Y_t &= \beta_{0t} + \beta_{1t} X_t + \nu_t, \qquad \nu_t \sim N[0, V_t] \\ \beta_{0t} &= \beta_{0(t-1)} + \omega_{\beta_{0t}}, \\ \beta_{1t} &= \beta_{1(t-1)} + \omega_{\beta_{1t}}, \end{aligned} \tag{1.50}$$

where $\boldsymbol{\omega}_t = (\omega_{t1}, \omega_{t2})' \sim N[0, \mathbf{W}_t]$. Then the Dynamic Linear Regression Models (DLRM) is defined by the quadruple $\{\mathbf{F}_t, \mathbf{I}, V_t, \mathbf{W}_t\}$.

The usual linear regression model formulated by the classical statistical theory corresponds to the static DLM which has $\boldsymbol{\theta}_t = \boldsymbol{\theta}$ and $V_t = V$ as known and constant in time. The system equation is redundant as the state parameter vector $\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1}$ is constant. The classical linear regression model assumes that there is a unique static vector of parameters for all the time points. See Soudant et al. (1997) for an example where DLRM is employed to show the relevance of time-varying parameter values in phytoplankton ecology.

1.5.2 General Multivariate Dynamic Linear Models

The univariate DLMs can be easily extended in an obvious way to multivariate problems. West and Harrison (1997) show that this can be done simply by taking the observations at each time t as vectors rather than scalars. In their extension, the observational errors are vectors, so it is required that observational variance matrices be defined for the joint stochastic structure of the observations conditional on the state parameters. This univariate theory extends nicely when it is assumed that the observational error variance matrices are known for all time t. West and Harrison (1997) provide theory for models with known variance matrices.

The distribution of the underlying time series conditional on the regression parameters and the distribution of the parameters conditional on the previous parameter state can be characterized by a set of quadruples

$$\{\mathbf{F}, \mathbf{G}, \mathbf{V}, \mathbf{W}\}_t = \{\mathbf{F}_t, \mathbf{G}_t, \mathbf{V}_t, \mathbf{W}_t\} \tag{1.51}$$

for each time t, where West and Harrison (1997) define for the corresponding model equations for each t as:

$$\begin{aligned} \mathbf{Y}_t &= \mathbf{F}_t' \boldsymbol{\theta}_t + \boldsymbol{\nu}_t, \qquad \boldsymbol{\nu}_t \sim N[\mathbf{0}, \mathbf{V}_t] \\ \boldsymbol{\theta}_t &= \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \qquad \boldsymbol{\omega}_t \sim N[\mathbf{0}, \mathbf{W}_t] \end{aligned} \tag{1.52}$$

where the error sequences $\boldsymbol{\nu}_t$ and $\boldsymbol{\omega}_t$ are independent and mutually independent. As usual in univariate DLM's, $\boldsymbol{\theta}_t$ is the n-dimensional state vector. The updating equation follow those

of the univariate case. For more detail on the General Multivariate DLM refer to West and Harrison (1997) pp. 582 - 584.

1.6 Simulation Techniques

In the Dynamic Linear Models previously introduced, the system and observation matrices $\mathbf{G}_t$ and $\mathbf{F}_t$ are set to specific values as part of the model specification. The only possible unknown parameters are those that are part of the variance matrices $\mathbf{W}_t$ and $\mathbf{V}_t$. In the following section we will discuss several popular ways deriving marginal densities or joint densities of the parameters of interest.

1.6.1 Metropolis-Hasting Algorithm

A very general method for constructing a Markov chain is the *Metropolis-Hastings algorithm* (Metropolis et al. 1953; Hastings 1970). The Metropolis-Hastings algorithm is a means of simulating a random sample from a probability distribution. In this procedure, we perform a random walk over the sample space, where each step consists of a proposal of a new state and a subsequent acceptance or rejection of the proposed state according to probabilities defined by the desired distribution, $p(\theta|\mathbf{y})$. We stop the process after a sufficient number of steps (or burn-in period), and observe the current state. In the limit, this observation will be a sample from the desired distribution, $p(\theta|\mathbf{y})$.

The method begins at $t = 0$ with the selection of $\theta^{(0)}$ drawn at random from some starting distribution $g(\theta)$, with the requirement that $p(\theta^{(0)}|\mathbf{y}) > 0$. Given $\theta^{(t)}$, the algorithm generates $\theta^{(t+1)}$ as follows:

1. Sample a candidate value θ^* from a proposal distribution $g(\cdot|\theta^{(t)})$.
2. Compute the *Metropolis-Hasting* acceptance ratio r, where
$$r = \frac{p(\theta^*|\mathbf{y})/g(\theta^*|\theta^{(t-1)})}{p(\theta^{(t-1)}|\mathbf{y})/g(\theta^{(t-1)}|\theta^*)} \tag{1.53}$$
3. Set
$$\theta^{(t)} = \begin{cases} \theta^* & \text{with probability } \min(1,r) \\ \theta^{(t-1)} & \text{otherwise.} \end{cases} \tag{1.54}$$
4. Increment t and return to step 1.

1.6.2 Gibbs Sampling

Gelman et al. (2004) state that a particular Markov chain algorithm that has been found useful in many multidimensional problems is the *Gibbs sampler*. The Gibbs Sampling algorithm was first presented by Geman and Geman (1984). The goal is to construct a Markov chain whose stationary distribution or some marginalization thereof equals the target distribution, $p(\boldsymbol{\theta}|\mathbf{y})$. The Gibbs sampler does this by sequentially sampling the d subvectors of the parameter vector $\boldsymbol{\theta}$, where $\boldsymbol{\theta} = (\theta_1, \ldots \theta_d)$. Each iteration of the Gibbs sampler cycles through the sub vectors of θ, drawing each subset conditional on the value of all the others. There are thus d steps in iteration t. At each iteration t, an ordering of the d subvectors of θ is chosen and, in turn, each θ_j^t is sampled from the conditional distribution given all the components of $\boldsymbol{\theta}$ such as follows:

$$\begin{aligned}
&\text{draw } \theta_1^{t+1} \text{ from } h(\theta_1|\theta_2^t, \ldots, \theta_d^t) \\
&\text{draw } \theta_2^{t+1} \text{ from } h(\theta_2|\theta_1^{t+1}, \ldots, \theta_3^t, \ldots, \theta_d^t) \\
&\qquad \vdots \\
&\text{draw } \theta_d^{t+1} \text{ from } h(\theta_d|\theta_1^{t+1}, \ldots, \theta_{d-1}^{t+1}).
\end{aligned}$$

Thus, each subvector θ_j is updated conditional on the latest value of θ for the other components, which are the iteration t values for the components already updated and the iteration $t-1$ values for the others. For more on the Gibbs Sampler, see Casella and George (1992); Gelfand and Smith (1990); Gelfand (2000).

1.6.3 Sampling Importance Resampling

Sampling importance resampling (SIR) is an algorithm that simulates samples approximately from some target distribution. Sampling importance resampling is based on the idea of importance sampling. Importance sampling draws values indirectly from a *target* distribution , f, by making use of an *proposal* distribution, g. Each point in the sample is weighted to correct the sampling probabilities so that the weighted sample can be related to the target distribution. For a detailed discussion on importance sampling refer to Rubin (1987), Rubin (1988), Smith and Gelfand (1992), Gelman et al. (2004), Givens and Hoeting (2005), and Albert (2007).

Givens and Hoeting (2005) describes the algorithm in the general case by letting $\boldsymbol{\Theta} = (\theta_1, \theta_2, \ldots, \theta_p)$ denote a random vector with density $f(\boldsymbol{\theta})$, and $g(\boldsymbol{\theta})$ denote the density corresponding to a multivariate proposal for f. For the target distribution f, the weights used to correct sampling probabilities are defined as

$$w(\boldsymbol{\theta}_i) = f(\boldsymbol{\theta}_i)/g(\boldsymbol{\theta}_i). \tag{1.55}$$

In the SIR algorithm these weights are converted to probabilities by the formula,

$$p(\boldsymbol{\theta}_i) = \frac{w(\boldsymbol{\theta}_i)}{\sum_{j=1}^{m} w(\boldsymbol{\theta}_i)}. \tag{1.56}$$

The probabilities in equation (1.56) are called the *standardized importance weights* for collection of values $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_m$ drawn independently identically distributed from the proposal distribution $g(\boldsymbol{\theta})$.

The SIR algorithm therefore proceeds as follows:

1. Sample candidates $\boldsymbol{\theta}_1^*, \ldots, \boldsymbol{\theta}_m^*$ i.i.d. from $g(\boldsymbol{\theta})$;
2. Calculate the standardized importance weights, $w(\boldsymbol{\theta}_1^*), \ldots, w(\boldsymbol{\theta}_m^*)$;
3. Resample $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n$ from $\boldsymbol{\theta}_1^*, \ldots, \boldsymbol{\theta}_m^*$ with replacement with probabilities $w(\boldsymbol{\theta}_1^*), \ldots, w(\boldsymbol{\theta}_m^*)$.

A random variable $\boldsymbol{\Theta}$ drawn with the SIR algorithm has distribution that converges to f as $m \to \infty$. It is noted that it is important to consider the relative sizes of the initial sample and the resample. These samples are m and n, respectively. In principle, it is required that $n/m \to 0$ for distributional convergence.

1.7 References

[1] Aitchison, J., and Dunsmore, I. (1975). *Statistical Prediction Analysis*. Cambridge University Press.

[2] Albert, J. (2007). *Bayesian Computation with R* (Vol. 747389981). New York: Springer.

[3] Berkson, J. (1969). Estimation of a linear function for a calibration line; consideration of a recent proposal. *Technometrics*. **11**, 649-660.

[4] Berger, J. O., and Bernardo, J. M. (1992). On the Development of Reference Priors. *Bayesian Statistics*. **4**(4), 35-60.

[5] Bernardo, J. M. (1979). Reference Posterior Distributions for Bayesian Inference. *Journal of the Royal Statistical Society. Series B (Methodological)*. 113-147.

[6] Bernardo, J. M., and Smith, A. F. M. (1994). *Bayesian Theory*. John Wiley and Sons.

[7] Brown, P. J. (1982). Multivariate calibration. *Journal of the Royal Statistical Society B*. **44**, 287-321.

[8] Brown, P. J. (1993). *Measurement, Regression, and Calibration*. Oxford University Press.

[9] Brown, P. J. and Sundberg, R. (1987). Confidence and conflict in multivariate calibration. *Journal of the Royal Statistical Society B*. **49**, 46-57.

[10] Casella, G., and George, E. I. (1992). Explaining the Gibbs Sampler. *The American Statistician*, **46**(3), 167-174.

[11] Clark, R. M. (1979). Calibration, Cross-validation and Carbon-14. **I**. *Journal of the Royal Statistical Society. Series A*. **142**, 47-62.

[12] Clark, R.M. (1980). Calibration, Cross-validation and Carbon-14. **II**. *Journal of the Royal Statistical Society. Series A*. **143**, 177-194.

[13] du Plessis, J. L., and van der Merwe, A. J. (1995). A Bayesian Approach to Multivariate and Conditional Calibration. *Computational Statistics & Data Analysis*. **19**(5), 539-552.

[14] Eisenhart, C. (1939). The interpretation of certain regression methods and their use in biological and industrial research. *Annals of Mathematical Statistics*. **10**, 162-186.

[15] Eno, D. R. (1999). Noninformative Prior Bayesian Analysis for Statistical Calibration Problems. Doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

[16] Easterling, R. G. (1981). [A Bayesian Analysis of the Linear Calibration Problem]: Discussion. *Technometrics*. **23**, 343-344.

[17] Fearn, T. (1983). Misuses of Ridge Regression in the Calibration of a Near Infra-red Reflectance Instrument. *Appl. Statist.* **32**, 73-79.

[18] Fieller, E. C. (1954). Some problems in interval estimation. *Journal of the Royal Statistical Society B*. **16**, 175-185.

[19] Fox, D. R. (1989). Conditional Multivariate Calibration. *Communications in Statistics-Theory and Methods*. **18**(6), 2311-2330.

[20] Gelfand, A. E., and Smith, A. F. (1990). Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association*, **85**(410), 398-409.

[21] Gelfand, A. E. (2000). Gibbs Sampling. *Journal of the American Statistical Association*, **95**(452), 1300-1304.

[22] Gelman, A. B., Carlin, J.S., Stern, H.S., and Rubin, D.B. (2004). *Bayesian Data Analysis*. Chapman and Hall/CRC.

[23] Geman, S., and Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on.* (6), 721-741.

[24] Givens, G. H., and Hoeting, J. A. (2005). *Computational Statistics.* John Wiley and Sons.

[25] Gruet, M. (1996). A Nonparametric Calibration Analysis. *The Annals of Statistics.* **24**(4), 1474-1492.

[26] Halperin, M. (1970). On Inverse Estimation in Linear Regression. *Technometrics.* **12**, 727-736.

[27] Harrison, P. J. and Stevens, C. F. (1976). Bayesian Forecasting. *Royal Statistical Society (B).* **38**, 205-247.

[28] Hastings, W. K. (1970). Monte Carlo Sampling Methods using Markov Chains and their Applications. *Biometrika.* **57**(1), 97-109.

[29] Hickey, G. L. (2006). The Linear Calibration Problem: A Bayesian Analysis. Doctoral dissertation, University of Durham, Durham, England.

[30] Hill, B. M. (1981). [A Bayesian Analysis of the Linear Calibration Problem]: Discussion. *Technometrics.* **23**, 335-338.

[31] Hoadley, B. (1970). A Bayesian look at Inverse Linear Regression. *Journal of the American Statistical Association.* **65**, 356-369.

[32] Hunter, W., and Lamboy, W. (1981). A Bayesian Analysis of the Linear Calibration Problem. *Technometrics.* **23**, 323-328.

[33] Jeffreys, H. (1946). An Invariant Form for the Prior Probability in Estimation Problems. *Proceedings of the Royal Statistical Society of London, Series A.* **186**, 453-461.

[34] Kalbfleisch, J. D. and Sprott, D. A. (1970). Applications of Likelihood Methods to Models Involving Large Numbers of Parameters (with discussion). *J. R. Statist. Soc. B* **32**, 175-208.

[35] Kalman, R. E. (1960), A New Approach to Linear Filtering and Prediction Problems, *Journal of Basic Engineering.* **83**, 35-45.

[36] Knafl, G., Spiegelman, C., Sacks, J. and Ylvisaker, D. (1984). Nonparametric Calibration. *Technometric.* **26**, 233-241.

[37] Krutchkoff, R. G. (1967). Classical and Inverse Regression Methods of Calibration. *Technometrics.* **9**, 425-439.

[38] Lawless, J. F. (1981). [A Bayesian Analysis of the Linear Calibration Problem]: Discussion. *Technometrics*. **23**, 334-335.

[39] Lechner, J., Reeve, C. and Spiegelman, C. (1982). An Implementation of the Scheffé Approach to Calibration using Spline Functions, Illustrated by a Pressure-Volume Calibration. *Technometric*. **2**(4), 229-234.

[40] Lwin, T. (1981). [A Bayesian Analysis of the Linear Calibration Problem]: Discussion. *Technometrics*. **23**, 339-341.

[41] Lwin, T. and Maritz, J. S. (1980). A Note on the Problem of Statistical Calibration. *Appl. Statist.* **29**, 135-141.

[42] Lwin, T. and Maritz, J. S. (1982). An Analysis of the Linear Calibration Controversy from the Perspective of Compound Estimation. *Technometrics*. **24**, 235-242.

[43] Lwin, T. and Spiegelman, C. H. (1986). Calibration with Working Standards. *Applied Statistics*. **35**. 256-261.

[44] Madansky, A. (1959). The Fitting of Straight Lines When both Variables are Subject to Error, *Journal of the American Statistical Association*. **54**(285), 173-205.

[45] Martinelle, S., (1970). On the Choice of Regression in Linear Calibration. *Technometrics*. **12**(1), 157-161.

[46] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of State Calculations by Fast Computing Machines. *Journal of Chemical Physics*. **21**, 1087.

[47] Misquitta, P. and Ruymgaart, F. H. (2005). Some Results on Nonparametric Calibration. *Communications in Statistics - Theory and Methods*. **34**(7), 1605-1616.

[48] Moore, D., McCabe, G. P., and Craig, B. (2010). *Introduction to the Practice of Statistics*. 7th Edition. Freeman, W. H. & Company.

[49] Naes, T., Irgens, C. and Martens, H. (1986). Comparison of Linear Statistical Methods for Calibration of NIR Instruments. *Appl. Statist.* **35**,195-206.

[50] Orban, J. E. (1981). [A Bayesian Analysis of the Linear Calibration Problem]: Discussion. *Technometrics*. **23**, 342-343.

[51] Osborne, C. (1990). Nonparametric Calibration. Doctoral dissertation, University of Bath.

[52] Osborne, C. (1991). Statistical Calibration: A Review. *International Statistical Review*. **59**, 309-336.

[53] Özyurt, Ö., and Erar, A. (2003). Multivariate Calibration in Linear Regression and its Application. *Hacettepe Journal of Mathematics and Statistics*. **32**, 53-63.

[54] Pole, A., West, M., and Harrison, P. J. (1994). *Applied Bayesian Forecasting and Time Series Analysis*. Chapman and Hall. New York.

[55] Rosenblatt, J. R., and Spiegelman, C. H. (1981). [A Bayesian Analysis of the Linear Calibration Problem]: Discussion. *Technometrics*. **23**, 329-333.

[56] Rubin, D. B. (1987). The Calculation of Posterior Distributions by Data Augmentation: Comment: A Noniterative Sampling/Importance Resampling Alternative to the Data Augmentation Algorithm for Creating a Few Imputations when Fractions of Missing Information are Modest: The SIR Algorithm. *Journal of the American Statistical Association*, **82**(398), 543-546.

[57] Rubin, D. B. (1988). Using the SIR Algorithm to Simulate Posterior Distributions. in *Bayesian Statistics 3*, eds. J. M. Bernardo, M. H. DeGroot, D. V. Lindley, and A. F. M. Smith, Cambridge, MA: Oxford University Press, 395-402.

[58] Scheffé, H. (1973). A Statistical Theory of Calibration. *Ann. Statist*. **1**, 1-37.

[59] Shukla, G. K. (1972). On the Problem of Calibration. *Technometrics*. **14**(3), 547-553.

[60] Smith, R. L., and Corbett, M. (1987). Measuring marathon courses: an application of statistical calibration theory. *Applied Statistics*. 283-295.

[61] Smith, A., and Gelfand, A. (1992). Bayesian Statistics without Tears: A Sampling-Resampling Perspective. *The American Statistician*. **46**(2), 84-88.

[62] Soudant, D., B. Beliaeff1, and G. Thomas (1997), Explaining *Dinophysis* cf. Acuminata Abundance in Antifer (Normandy, France) using Dynamic Linear Regression, *Marine Ecology Progress Series*, **156**, 67-74.

[63] Sundberg, R. and Brown, P.J. (1989). Multivariate Calibration with More Variables than Observations. *Technometrics*. **31**, 365-371.

[64] Tibshirani, R .J. (1988). Estimating transformations for regression via additivity and variances stabilisation. *Journal of the American Statistical Association*. **83**, 394-405.

[65] Walker, E. L., Starnes, B. A., Birch, J. B., and Mays, J. E. (2010). Model Robust Calibration. Unpublished paper presented at 27_{th} AIAA Aerodynamic Measurement Technology and Ground Testing Conference, Chicago, Illinois

[66] West, M., Harrison, P. J., and Migon, H. S. (1985). Dynamic generalized linear models and Bayesian forecasting. *Journal of the American Statistical Association*, **80**(389), 73-83.

[67] West, M. and P.J. Harrison (1997). Bayesian Forecasting and Dynamic Linear Models, 2nd Edition. Springer-Verlag, New York.

[68] Williams, E. J. (1969a). Regression Methods in Calibration Problems. *Bull. Int. Statistics*. **43**, 17-28.

[69] Williams, E. J. (1969b). A Note on Regression Methods in Calibration. *Technometrics*. **11**(1), 189-192.

Chapter 2

A Dynamic Approach to Linear Statistical Calibration with an Application in Microwave Radiometry

Abstract

The problem of statistical calibration of a measuring instrument can be framed both in a statistical context as well as in an engineering context. In the first, the problem is dealt with by distinguishing between the "classical" approach and the "inverse" regression approach. Both of these models are static models and are used to estimate "exact" measurements from measurements that are affected by error. In the engineering context, the variables of interest are considered to be taken at the time at which you observe them. The Bayesian time series analysis method of Dynamic Linear Models (DLM) can be used to monitor the evolution of the measures, thus introducing a *dynamic* approach to statistical calibration. The research presented employs the use of Bayesian methodology to perform statistical calibration. The DLM framework is used to capture the time-varying parameters that may be changing or drifting over time. Two separate DLM based models are presented in this paper. A simulation study is conducted where the two models are compared to some well known 'static' calibration approaches in the literature from both the frequentist and Bayesian perspectives. The focus of the study is to understand how well the *dynamic statistical calibration* methods perform under various variance ratios, r. The posterior distributions of the estimated calibration points as well as the 95% coverage intervals are compared by statistical summaries. These dynamic methods are applied to a microwave radiometry dataset.

2.1 Introduction

Calibrating measurement instruments is a important problem that engineers frequently need to address. There exist several statistical methods that address this problem that are based on a simple linear regression approach. In traditional simple linear regression, the goal is to relate a known value of X to an uncertain value of Y using a linear relationship. In contrast, the statistical calibration problem seeks to utilize a simple linear regression model to relate a known value of Y to an uncertain value of X. This is why statistical calibration is sometimes called *inverse regression* due to its relationship to simple linear regression (Osborne 1991; Ott and Longnecker 2009). Recall in linear regression the model is given as follows:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{2.1}$$

where $\mathbf{Y}$ is a $(n \times 1)$ response vector, $\mathbf{X}$ is a $(n \times p)$ matrix of independent variables with $p = k+1$ total model parameters, $\boldsymbol{\beta}$ is a $(p \times 1)$ vector of unknown fixed parameters and $\boldsymbol{\epsilon}$ is a $(n \times 1)$ vector of uncorrelated error terms with zero mean (Myers 1990; Draper and Smith 1998; Montgomery *et al.* 2012). It is assumed that the values of the predictor variable $\mathbf{X} = \mathbf{x}$ are nonrandom and observed with negligible error, while the n error terms are random variables with mean zero and constant variance σ^2 (Myers 1990). Typically, in regression, of interest is the estimation of the parameter vector, $\boldsymbol{\beta}$, and possibly the prediction of a future value $\hat{\mathbf{Y}}_{i|new}$ corresponding to a new $\mathbf{X} = x'_{i|new}$ value. The prediction problem is relatively straightforward, due to the fact that a future $\mathbf{Y}_i$ value can be made directly by substituting $x'_{i|new}$ into (2.1) with $E[\epsilon] = 0$.

For the statistical calibration problem, let y_0 be the known *observed* value of the response and x_0 be the corresponding regressor that is to be estimated. This problem is conducted in two stages: first measurement pairs (x_i, y_i) of data are observed and a simple linear regression line is fit by estimating $\boldsymbol{\beta}$; secondly, m observations of the response are observed, all corresponding to a single x_0 (Özyurt and Erar 2003). Since y_0 is fixed, inferences are different than those in a traditional regression (or prediction) problem (Osborne 1991; Eno 1999; Eno and Ye 2000).

2.1.1 Classical Calibration Methods

Eisenhart (1939) offered the first solution to the calibration problem, and is commonly known as the "*classical*" estimator to the linear calibration problem. He assumed that the relationship between x and y was of a simple linear form:

$$E(Y|X = x) = \beta_0 + \beta_1 x.$$

The estimated regression line for the first stage of the experiment is given by

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X, \tag{2.2}$$

where $\hat{\beta}_0$ and $\hat{\beta}_1$ are the least squares estimates of β_0 and β_1, respectively. Using the data collected at the first stage of experimentation, Eisenhart (1939) inverts Equation (2.2) to estimate the unknown regressor value x_0 for an observed response value y_0, by:

$$\hat{x}_{0,c} = \frac{y_0 - \hat{\beta}_0}{\hat{\beta}_1} \tag{2.3}$$

where $\hat{x}_{0,c}$ denotes the "*classical*" estimator for x_0. Since division by $\hat{\beta}_1$ is used there is an implicit assumption that $|\hat{\beta}_1| > 0$.

Assuming that $|\hat{\beta}_1| > 0$, Brown (1993) describes the following interval estimate corresponding to Eisenhart (1939):

$$\frac{y_0 - \hat{\beta}_0}{\hat{\beta}_1}\left(1 + \frac{\hat{\sigma}^2 t^2_{\alpha/2}}{\hat{\beta}_1^2 S_{xx}}\right) \pm \frac{\hat{\sigma} t_{\alpha/2}}{\hat{\beta}_1}\left(1 + \frac{1}{2n} + \frac{(y_0 - \hat{\beta}_0)^2 + \hat{\sigma}^2 t^2_{\alpha/2}}{2\hat{\beta}_1^2 S_{xx}}\right),$$

where

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2}{n-2}},$$

$$S_{xx} = \sum_{i=1}^{n}(x - \bar{x})^2,$$

and $t_{\alpha/2}$ is a critical t-value with $n-2$ degrees of freedom.

Krutchkoff (1967) proposed a competitive approach to Eisenhart's (1939) classical linear calibration solution, which he called the "*inverse*" regression calibration method and is written as:

$$X_i = \phi + \delta Y_i + \epsilon_i^{'},$$

where ϕ and δ are the parameters in the linear relationship and $\epsilon_i^{'}$ are independent identically distributed measurement errors with a zero mean and finite variance. Here ϕ and δ are estimated via least squares. The unknown x_0 can be estimated directly by substituting y_0 into the fitted equation:

$$\hat{x}_{0,I} = \hat{\phi} + \hat{\delta} y_0. \tag{2.4}$$

We let $\hat{x}_{0,I}$ denote the "*inverse*" estimator of x_0. The $100(1-\alpha)\%$ confidence interval for $E(x_{0,I}|y_0)$ can be written as

$$x_{0,I}(y_0) \pm t_{\alpha/2}\hat{\sigma}\sqrt{\frac{1}{n} + \frac{(y_0 - \bar{y})^2}{S_{yy}}}$$

where

$$S_{yy} = \sum_{i=1}^{n}(y_i - \bar{y})^2.$$

Krutchkoff (1967) used a simulation study, where he found that the mean squared error of estimation for x_0 was uniformly less for this estimator versus the classical estimator. The inverse approach was later supported by Lwin and Maritz (1982). For criticisms of Krutchkoff's (1967) approach, such as bias, see Osborne (1991).

2.1.2 Bayesian Calibration Methods

The first noted Bayesian solution to the calibration problem was presented by Hoadley (1970). His work was motivated by the unanswered question in the Frequentist community of whether β_1 is zero (or close to zero). Hoadley (1970) justified the use of the "*inverse*" estimator (Krutchkoff, 1967) by considering the usual F-statistic to test the hypothesis that $\beta_1 = 0$ where $F = \hat{\beta}_1^2 S_{xx}/\hat{\sigma}^2$,

$$\hat{\sigma}^2 = \frac{\left\{\sum_{i=1}^{n}\left(y_{1i} - (\hat{\beta}_0 + \hat{\beta}_1 x_i)\right)^2 + \sum_{j=1}^{m}(y_{2j} - \bar{y}_2)^2\right\}}{(n+m-3)}.$$

The assumption made by Hoadley (1970) reflects that x_0 is random and *a priori* independent of $\pi(\beta_0, \beta_1, \sigma^2)$, so that the joint prior distribution of $\pi(\beta_0, \beta_1, \sigma^2, x_0) \propto \pi(\beta_0, \beta_1, \sigma^2)\pi(x_0)$. Hoadley (1970) first assumed that $(\beta_0, \beta_1, \sigma^2)$ had a uniform distribution,

$$\pi(\beta_0, \beta_1, \sigma^2) \propto \sigma^{-2},$$

but the prior distribution for x_0 was not given.

Hoadley (1970) shows for $m = 1$ (one observation at the prediction stage), that if x_0 has a prior density from a Student t distribution with $n-3$ degrees of freedom, a mean of 0, and a scale parameter

$$\sigma = \frac{n+1}{n-3},$$

the posterior distribution is

$$\pi(x_0|\mathbf{Data}) = t_{n-2}\left(\hat{x}_{0,I}, \left[\frac{n+1+(\hat{x}_{0,I})^2/R}{F+n-2}\right]\right), \tag{2.5}$$

where $\hat{x}_{0,I}$ is the inverse estimator given by (2.4), $R = \frac{F}{F+n-2}$ and $F = \hat{\beta}_1^2 S_{xx}/\hat{\sigma}^2$.

Hunter and Lamboy (1981) also considered the calibration problem from a Bayesian point of view and their approach is similar to that of Hoadley (1970) because both assume the prior distribution to be

$$\pi(\beta_0, \beta_1, \sigma^2, \eta) \propto \sigma^{-2}$$

where $\eta = \beta_0 + \beta_1 x_0$ which is the predicted y_0. The primary difference between their approach and the approach of Hoadley (1970) is that *a priori* they assume that η and $(\beta_0, \beta_1, \sigma^2)$ are independent while Hoadley (1970) assumed *a priori* that x_0 and $(\beta_0, \beta_1, \sigma^2)$ are independent.

Hunter and Lamboy (1981) used an approximation to the posterior distribution of the unknown regressor x_0, given by

$$\pi(x_0|\mathbf{Data}) = N\left(\hat{x}_{0,c}, \frac{(s_{11}+s_{33})s_{22} - s_{12}^2}{s_{22}\hat{\beta}_1^2}\right), \tag{2.6}$$

where

$$\mathbf{S} = \{s_{i,j}\} = \begin{bmatrix} s_{11} & s_{12} & 0 \\ s_{12} & s_{22} & 0 \\ 0 & 0 & s_{33} \end{bmatrix} = \begin{bmatrix} (\mathbf{X}'\mathbf{X})^{-1}\hat{\sigma}^2 & \mathbf{0} \\ \mathbf{0} & \hat{\sigma}^2/m \end{bmatrix},$$

with $\hat{x}_{0,c}$ being the classical estimator given in Equation (2.3), $s_{i,j}$ denoting the element of the i^{th} row and j^{th} column from the variance-covariance matrix of the joint posterior density of (β_0, β_1, η).

The remainder of this chapter is organized as follows. Section 2.2 presents the development of the dynamic approaches to the statistical calibration problem. In Section 2.3, the results from the simulation study where the dynamics methods are evaluated along with the static approaches are presented. In Section 2.4, the proposed methods are applied to microwave radiometer data. In Section 2.5, future work and other considerations are given.

2.2 Dynamic Linear Calibration Approach

Traditional calibration methods assume the regression relationship is "static" in time. In many cases this is false. For example, in microwave radiometry the static nature of the relationship is known to change across time. A dynamic approach can be created by letting the regression coefficients vary through time,

$$y_t = \beta_{0t} + \beta_{1t}x_t + \epsilon_t,$$

where $\epsilon_t \overset{iid}{\sim} N[0, \sigma_t^2]$ and is known as the *observational* error.

The model may have different defining parameters at different times. One approach is to model β_{0t} and β_{1t} by using random walk type evolutions for the defining parameters, such as:

$$\begin{aligned} \beta_{0t} &= \beta_{0(t-1)} + \omega_{\beta_{0t}}, \\ \beta_{1t} &= \beta_{1(t-1)} + \omega_{\beta_{1t}}, \end{aligned}$$

where $\omega_{\beta_{0t}}$ and $\omega_{\beta_{1t}}$ are independent zero-mean error terms with finite variances. At any time t the calibration problem is given by:

$$y_{0t} = \beta_{0t} + \beta_{1t}x_{0t} + \epsilon_t, \qquad t = 1, 2, \ldots, T.$$

The Bayesian Dynamic Linear Models (DLMs) approach of West *et al.* (1985) and West and Harrison (1997) can be employed to achieve this goal. Recall the DLM framework is:

$$\begin{aligned} \text{Observation equation :} \quad & \mathbf{Y}_t = \mathbf{X}_t\boldsymbol{\theta}_t + \boldsymbol{\epsilon}_t, & \boldsymbol{\epsilon}_t \sim N_r[\mathbf{0}, \mathbf{E}] \\ \text{System equation :} \quad & \boldsymbol{\theta}_t = \mathbf{G}_t\boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, & \boldsymbol{\omega}_t \sim N_d[\mathbf{0}, \mathbf{W}] \\ \text{Initial information :} \quad & (\boldsymbol{\theta}_0|D_0) \sim N_d[\mathbf{m_0}, \mathbf{C_0}], & \end{aligned}$$

for some prior mean $\mathbf{m_0}$ and variance $\mathbf{C_0}$ with the vector of error terms, $\boldsymbol{\epsilon}_t$ and $\boldsymbol{\omega}_t$ independent across time and at any time.

To update the model through time, West and Harrison (1997) give the following method:

(a) Posterior distribution at $t-1$: For some mean $\mathbf{m}_{t-1}$ and variance $\mathbf{C}_{t-1}$,
$$(\boldsymbol{\theta}_{t-1}|D_{t-1}) \sim N_d[\mathbf{m}_{t-1}, \mathbf{C}_{t-1}].$$

(b) Prior distribution at time t: $(\boldsymbol{\theta}_t|D_{t-1}) \sim N_d[\mathbf{a}_t, \mathbf{R}_t]$, where
$$\mathbf{a}_t = \mathbf{G}_t\mathbf{m}_{t-1} \qquad \text{and} \qquad \mathbf{R}_t = \mathbf{G}_t\mathbf{C}_{t-1}\mathbf{G}_t^{'} + \mathbf{W}.$$

(c) One-step forecast: $(\mathbf{Y}_t|D_{t-1}) \sim N_r[\mathbf{f}_t, \mathbf{Q}_t]$, where
$$\mathbf{f}_t = \mathbf{X}_t\mathbf{a}_t \qquad \text{and} \qquad \mathbf{Q}_t = \mathbf{X}_t\mathbf{R}_t\mathbf{X}_t^{'} + \mathbf{E}.$$

(d) Posterior distribution at time t: $(\boldsymbol{\theta}_t|D_t) \sim N_d[\mathbf{m}_t, \mathbf{C}_t]$,
$$\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_t\mathbf{e}_t \qquad \text{and} \qquad \mathbf{C}_t = \mathbf{R}_t - \mathbf{A}_t^{'}\mathbf{Q}_t\mathbf{A}_t,$$
where
$$\mathbf{A}_t = \mathbf{Q}_t^{-1}\mathbf{X}_t\mathbf{R}_t \qquad \text{and} \qquad \mathbf{e}_t = \mathbf{Y}_t - \mathbf{f}_t.$$

The DLM framework is used to establish the evolving relationship between the fixed design matrix $\mathbf{X}_t$ and $\mathbf{Y}_t$ by estimating $\boldsymbol{\theta}_t$, which is a $(d \times n)$ matrix of time-varying regression coefficients β_{0t} and β_{1t}. For our calibration situation $\mathbf{Y}_t$ is a $(r \times n)$ matrix of responses and $\mathbf{G}_t$ is a known $(d \times d)$ system matrix. The error terms $\boldsymbol{\epsilon}_t$ and $\boldsymbol{\omega}_t$ are independent normally distributed random $(r \times n)$ matrices with zero mean and constant variance-covariance matrices $\mathbf{E}$ and $\mathbf{W}$. For simplification $\mathbf{G}_t$ is set equal to $\mathbf{I}_{(d\times d)}$, $\mathbf{E}$ is set equal to $\sigma_E^2\mathbf{I}_{(r\times r)}$ and $\mathbf{W}$ is $\sigma_W^2\left[\mathbf{X}'\mathbf{X}\right]^{-1}$. The past information is contained in the set D_0.

We specify a prior in the first stage of calibration for the unknown variances and derive an algorithm to draw from the posterior distribution of the unknown parameters,

$$\pi(\boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2|\mathbf{Y}_t) \propto \pi(\boldsymbol{\theta}_t|\sigma_E^2, \sigma_W^2, \mathbf{Y}_t)\pi(\sigma_E^2, \sigma_W^2|\mathbf{Y}_t).$$

The second stage of the calibration experiment consists of using the joint posterior distribution $\pi(\boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2|\mathbf{Y}_t)$ to derive $x_{0t}|\boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2$ for each draw of $\pi(\sigma_E^2, \sigma_W^2|\mathbf{Y}_t)$. The estimator for the parameter of interest, x_{0t}, is defined in a manner akin to Eisenhart (1939), Hunter and Lamboy (1981), and Eno (1999), where

$$x_{0t} = \frac{y_{0t} - \beta_{0t}}{\beta_{1t}}. \tag{2.7}$$

In the final stage of the calibration experiment, the posterior distribution summary statistics are gathered at each time point t. The posterior median and credible intervals are taken for each t across the draws of $x_{0t}|\boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2$. The result of the dynamic calibration experiment is a time series of calibration distributions across time. We will be able to observe the distributional changes of the system with respect to the calibration reference.

The proposed calibration estimator is developed by first considering the joint posterior

distribution $\pi(\sigma_E^2, \sigma_W^2|\mathbf{Y}_t)$. We let $\mathbf{\Gamma}$ denote the vector of unknown DLM dispersion parameters where $\mathbf{\Gamma}' = (\sigma_E^2, \sigma_W^2)$. The prior information for the dispersion parameters is described by a prior density $\pi(\mathbf{\Gamma})$ which summarizes what is known about the variance parameters before any data are observed. Using the Bayesian inferential approach, the prior information about the parameters must be combined with information contained in the data. The information provided by the data is captured by the likelihood functions, $f_{\mathbf{Y}}(\mathbf{Y}_t|\boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2)$ and $f_{\boldsymbol{\theta}}(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}, \sigma_W^2)$ for the observation equation and the system equation, respectively. The combined information is described by the posterior density using the Bayes' theorem (Bernardo and Smith 1994) as

$$\pi(\mathbf{\Gamma}|\mathbf{Y}_t) \propto f_{\mathbf{Y}}(\mathbf{Y}_t|\boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2) \cdot f_{\boldsymbol{\theta}}(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}, \sigma_W^2) \cdot \pi(\mathbf{\Gamma}).$$

For our calibration problem it is believed that $\sigma_E^2 > \sigma_W^2$. To deal with the variance relationship we specify the following prior distributions:

$$\sigma_E^2 \sim Uniform(0, \alpha_E) \tag{2.8}$$
$$\sigma_W^2|\sigma_E^2 \sim Uniform(0, \sigma_E^2). \tag{2.9}$$

Prior distributions (2.8) and (2.9) ensures the system variance to be less than the observation variance. Since these are proper prior distributions the resulting posterior distribution will also be proper.

In the first stage of calibration, the joint distribution of the observations, states, and unknown parameters is as follows:

$$\begin{aligned}\pi(\mathbf{Y}_{1:T}, \boldsymbol{\theta}_{0:T}, \sigma_E^2, \sigma_W^2) &= f_{\mathbf{Y}}(\mathbf{Y}_{1:T}|\boldsymbol{\theta}_{0:T}, \sigma_E^2, \sigma_W^2) \cdot f_{\boldsymbol{\theta}}(\boldsymbol{\theta}_{0:T}|\sigma_W^2) \cdot \pi(\mathbf{\Gamma}) \\ &= \prod_{t=1}^{T} f_{\mathbf{Y}}(\mathbf{Y}_t|\boldsymbol{\theta}_t, \sigma_E^2) \cdot \prod_{t=1}^{T} f_{\boldsymbol{\theta}}(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}, \sigma_W^2) \\ &\quad \cdot \pi(\boldsymbol{\theta}_0) \cdot \pi(\sigma_E^2) \cdot \pi(\sigma_W^2|\sigma_E^2).\end{aligned}$$

where the likelihood for the observation equation is

$$f_{\mathbf{Y}}(\mathbf{Y}_t|\boldsymbol{\theta}_t, \sigma_E^2) \propto \sigma_E^{-T} \exp\left\{-\frac{1}{2\sigma_E^2}\sum_{t=1}^{T}(\mathbf{Y}_t - \mathbf{X}_t\boldsymbol{\theta}_t)^2\right\}$$

and the likelihood for the system equation is

$$f_{\boldsymbol{\theta}}(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}, \sigma_W^2) \propto \sigma_W^{-T} \exp\left\{-\frac{1}{2\sigma_W^2}\sum_{t=1}^{T}(\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t-1})^2\right\}.$$

Given the joint distribution above, the posterior distribution is

$$\pi(\mathbf{x}_{0t}|\boldsymbol{\theta}_t, \mathbf{\Gamma}, \mathbf{Y}_t) \tag{2.10}$$

where

$$\mathbf{x}_{0t} = \frac{\mathbf{y}^*_{\mathbf{0t}}}{\boldsymbol{\theta}_t} \tag{2.11}$$

and $\mathbf{y}^*_{0t} = \mathbf{y}_{0t} - \bar{y}_t$ ($\bar{y}_t$ is the cumulative mean of the observations up to time t) and $\boldsymbol{\theta}_t = \hat{\boldsymbol{\beta}}_{1t}$. Samples from the posterior distribution in Equation (2.10) are drawn by implementing the Sampling Importance Resampling (Albert 2007; Givens and Hoeting 2005) approach.

The development of the estimator in Equation (2.11) is deterministic in approach. We present a fully Bayesian approach to dynamic calibration that incorporates the uncertainty in estimation. The second dynamic calibration model is derived by Bayes' theorem

$$\pi(\mathbf{x_{0t}}|\mathbf{Y}_t) \propto \pi(\mathbf{x_{0t}}) f(\mathbf{Y}_t|\mathbf{x_{0t}}),$$

where $\pi(\mathbf{x_{0t}}|\mathbf{Y}_t)$ is the posterior distribution for $\mathbf{x}_{0t}$. The prior belief for the calibration values is denoted as $\pi(\mathbf{x_{0t}})$ with the $f(\mathbf{Y}_t|\mathbf{x_{0t}})$ denoting the likelihood function.

The objective of any Bayesian approach is to obtain the posterior distribution from which inferences can be made. Here the desired posterior is

$$\pi(\mathbf{x_{0t}}|\mathbf{Y}_t) \tag{2.12}$$

which must be dynamic through time. We determine the posterior distribution (2.12) in a similar manner as described above in Equations (2.10) and (2.11). In the first stage of the calibration experiment, the data is scaled and centered, therefore setting the $y-$intercept equal to zero and the reference measurements centered at zero. Centering of the data is used to reduce the parameter space. The posterior distribution can be thought of as:

$$\pi(\mathbf{z_{0t}}|\mathbf{Y}^*_t) \propto \pi(\mathbf{z_{0t}}) f(\mathbf{Y}^*_t|\mathbf{z_{0t}}), \tag{2.13}$$

with $\mathbf{z_{0t}}$ representing the transformed calibrated value at time t and $\mathbf{Y}^*_t = \mathbf{Y}_t - \bar{Y}_t$, where $\bar{Y}_t$ is the cumulative mean of the observations. Given this information *a priori* we define the prior distribution

$$\pi(\mathbf{z_{0t}}) = N(0, 1).$$

The posterior density in Equation (2.13) is defined as

$$\pi(\mathbf{z_{0t}}|\mathbf{Y}^*_t) \propto \exp\left\{-\frac{1}{2}\left[\sigma^{-2}_{Y_t}\sum_{t=1}^{T}(\boldsymbol{\xi}_t - \mathbf{z}_{0t})^2 + \mathbf{z}^2_{0t}\right]\right\} \tag{2.14}$$

where $\boldsymbol{\xi}_t = {}^{\mathbf{Y}^*_{0t}}/_{\boldsymbol{\theta}_t}$. Applying Bayes' theorem and completing the square, the posterior distribution is

$$\pi(\mathbf{z_{0t}}|\mathbf{Y}^*_t) \sim N(\mu_{z_{0t}}, \sigma^2_{z_{0t}}), \tag{2.15}$$

with

$$\begin{aligned} \mu_{z_{0t}} &= \frac{\boldsymbol{\xi}_t}{1+\sigma^2_{Y_t}}, \\ \sigma^2_{z_{0t}} &= \frac{1}{1+\sigma^2_{Y_t}} \end{aligned}$$

and

$$\sigma^2_{Y_t} = \text{tr}(\mathbf{Q}_t).$$

where tr($\cdot$) denotes trace of the one-step forecast variance-covariance matrix. We derive the posterior in Equation (2.12) by drawing from Equation (2.15) and transforming the data back to the original scale as so:

$$\mathbf{x}_{0t} = \bar{X} + \mathbf{z}_{0t}\sigma_X, \tag{2.16}$$

where $\bar{X}$ is the mean of the reference measurements vector and σ_X is the standard deviation of the reference measurements vector.

The dynamic calibration algorithm is developed for both of the approaches using R (R Development Core Team, 2013) and is conducted as below.

Algorithm 1: Dynamic Calibration

---

1. Generate M proposal samples for (σ^2_E, σ^2_W) from $\pi(\sigma^2_E)$ and $\pi(\sigma^2_W|\sigma^2_E)$;

2. Calibration data are fit using the DLM framework for each of the M proposal samples $(\sigma^{2(m)}_E, \sigma^{2(m)}_W)$, with the prior moments for $(\boldsymbol{\theta}_0|D_0)$ as $\mathbf{m_0} = \mathbf{1_d}$ and $\mathbf{C_0} = \mathbf{100I_{(d\times d)}}$, where $\mathbf{1_d}$ is a $d-$dimensional vector of ones.

   a. Data are scaled and shifted such that $\sum_{i=1}^r x_i = 0$, $\frac{1}{n}\sum_{i=1}^r x_i^2 = 1$ and $y-$intercept $= 0$, where $y^*_t = y_t - \bar{y}_t$ for all t (i.e. $\bar{y}_t$ is the cumulative mean up to time t);

   b. Estimate $\boldsymbol{\theta}^{(m)}_t|\sigma^{2(m)}_E, \sigma^{2(m)}_W$ for the m^{th} proposal sample is calculated for all t;

   c. Estimate $x^{(m)}_{0t}|\boldsymbol{\theta}^{(m)}_t, \sigma^{2(m)}_E, \sigma^{2(m)}_W$ for the m^{th} proposal sample is calculated for all t, using either Equation (2.11) or drawing from Equation (2.15);

   d. Calculate log-likelihood density weights, $log[f(\boldsymbol{\Gamma}^{(m)})]$, for each $(\sigma^{2(m)}_E, \sigma^{2(m)}_W)$ pair

3. Sampling Importance Resampling (SIR) is used to simulate samples of $x_{0t}|\boldsymbol{\theta}_t, \sigma^2_E, \sigma^2_W$ by accepting a subset of $N = 1,000$ from the proposal density to be distributed according to the posterior density $\pi(\boldsymbol{\Gamma}|\mathbf{Y}_t)$ with candidate density $\pi(\boldsymbol{\Gamma})$.

   a. Calculate the standardized importance weights, $w(\boldsymbol{\Gamma}^{(1)}), \ldots, w(\boldsymbol{\Gamma}^{(M)})$, where $w(\boldsymbol{\Gamma}^{(m)}) = log[f(\boldsymbol{\Gamma}^{(m)})] - log[g(\boldsymbol{\Gamma}^{(m)})]$ for the m^{th} proposal sample;

   b. Sample N calibrated time series from the M proposal values with replacement given probabilities $p(\boldsymbol{\Gamma}^{(m)})$ where

$$p(\boldsymbol{\Gamma}^{(m)}) = \frac{e^{w(\boldsymbol{\Gamma}^{(m)})}}{\sum_{j=1}^M e^{w(\boldsymbol{\Gamma}^{(j)})}}.$$

4. Rescale calibrated time series to original scale by Equation (2.16) and take summary statistics (i.e. medians and credible sets) across each time t .

2.3 Simulation Study

A simulation study, mirroring the microwave radiometer example in Section 2.4, considers the performance of the proposed dynamic calibration approaches to the static approaches discussed in Section 2.1. For notation, the calibration methods are labelled as follows:

1. M_{D1} is the first deterministic dynamic calibration model given in Equation (2.11);
2. M_{D2} is the Bayesian dynamic calibration model given by Equation (2.15);
3. M_{F1} is the "*classical*" approach of Eisenhart (1939) defined in Equation (2.3);
4. M_{F2} is the "*inverse*" approach of Krutchkoff (1967) defined in Equation (2.4);
5. M_{B1} is the Hoadley (1970) Bayesian approach as defined in Equation (2.5);
6. M_{B2} is the Hunter & Lamboy (1981) Bayesian approach as defined in Equation (2.6).

Note that static methods M_{F1}, M_{F2}, M_{B1}, and M_{B2} require that model fitting and the calibration take place after all the data has been collected. This is in contrast to the dynamic methods that both fit the model and generate calibrated values at each point through time and hence provide a near real time calibration. In order to assess the performance of the calibration methods 100 datasets were randomly generated according to

$$\mathbf{Y}_t = \mathbf{X}\boldsymbol{\theta}_t + \boldsymbol{\epsilon}_t, \tag{2.17}$$

where $\mathbf{X}$ is a known fixed design matrix of reference values. The number of reference measurements used in the study was two and five. The reference values at the first stage of the simulation study were equally spaced, covering the interval $[20, 100]$. For the two reference case, the fixed design matrix is

$$\mathbf{X} = \begin{bmatrix} 1 & 20 \\ 1 & 100 \end{bmatrix}$$

and for the five reference case the design matrix is

$$\mathbf{X} = \begin{bmatrix} 1 & 20 \\ 1 & 40 \\ 1 & 60 \\ 1 & 80 \\ 1 & 100 \end{bmatrix}.$$

The vector of regression parameters, $\boldsymbol{\theta}_t$, are randomly drawn from a multivariate normal distribution with mean vector $[12.7434\ 0.02655]'$ and variance-covariance matrix, $\boldsymbol{\Sigma} = \sigma_W^2 \left[\mathbf{X}'\mathbf{X}\right]^{-1}$ for $t = 1, \ldots, T$, where $T = 1000$. For each t, the random multivariate error vector is

$$\boldsymbol{\epsilon}_t \sim N_r[\mathbf{0}, \sigma_E^2 \mathbf{I}]$$

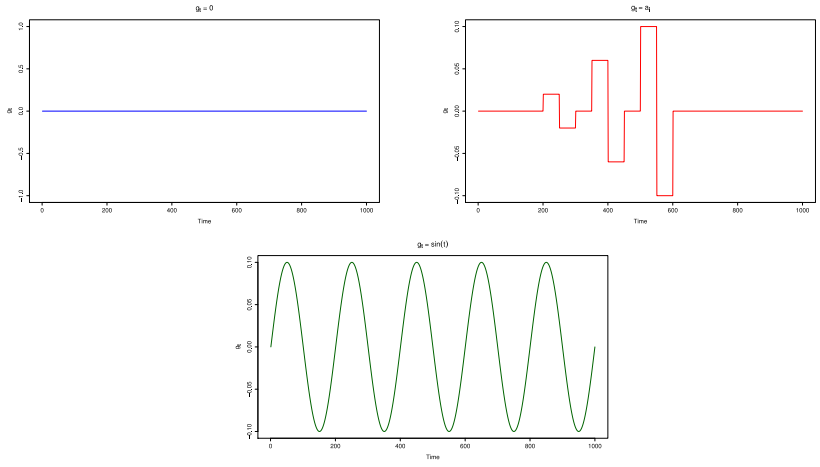


Figure 2.1: *Three distinct gain fluctuations: (a)* $g_t = 0$*; (b)* $g_t = a_i$ *with* $\sum_{i=1}^{n} a_i = 0$*; (c)* $g_t = 0.1sin(0.025t)$

where the errors are mutually independent. The relationship of the values for σ_E^2 and σ_W^2 will be explained later.

The dynamic and static calibration methods are evaluated for three distinct system fluctuations, g_t, on the regression slope calculated in the first stage of calibration. The value g_t is added to the slope component β_{1t} of $\boldsymbol{\theta}_t$, therefore making Equation (2.17)

$$y_{jt} = \beta_{0t} + (\beta_{1t} + g_t)x_j + \epsilon_t, \quad t = 1, \ldots, T$$

for the j^{th} calibration references. The three scenarios for the fluctuations g_t are as follows:

1. a constant zero ($g_t = 0$) for all t, representing a stable system;
2. a stable system with abrupt shifts ($g_t = a_i$) in system, with $\sum_{i=1}^{n} a_i = 0$; and
3. a constant sinusoidal fluctuation ($g_t = 0.1\text{sin}(0.025t)$) for all t.

Figure 2.1 illustrates the relationship of g_t across time.

The magnitude and relationship of the variance pair (σ_E^2, σ_W^2) influence the DLM and hence to study this influence we set the variances to reflect various *variance ratios*. The true values for σ_E^2 and σ_W^2 used in the simulation study are $(0.0001, 0.001, 0.01)$ and $(0.00001, 0.00005)$, respectively. Petris *et al.* (2009) define the variance ratio r as follows:

$$r = \frac{\text{Observation Variance}}{\text{System Variance}} = \frac{\sigma_E^2}{\sigma_W^2}.$$

The variance ratios r in the simulation study were examined in two sets. First, r is set equal to 10, 100, and 1000. Next, the ratio r was set to equal 2, 20, and 200. The variety of r values allows us to examine the methods under different levels of noise. Each simulation is

repeated 100 times for both the 2- and 5-point calibration models, thus providing us with 36 possible models for examination from the settings of r.

After the data was fit with each of the methods we considered the following measures for assessing the performance of the dynamic methods compared to the familiar static approaches: (1) average mean square error; (2) average coverage probability; and (3) average interval width. For each of the simulated data sets, the mean squared error (MSE) is calculated as

$$MSE = \frac{1}{T}\sum_{t=1}^{T}(\hat{x}_{0t} - x_{0t})^2.$$

The MSE are averaged across the 100 simulated data sets thus deriving an average mean squared error ($AvMSE$) as

$$AvMSE = \frac{1}{100}\sum_{j=1}^{100} MSE_j.$$

The coverage probability based on the 95% coverage interval is estimated for all of the calibration methods. The coverage interval for the dynamic and static Bayesian approaches is the 95% credible interval, and the 95% confidence interval is used for the frequentist methods. Note that for credible intervals x_{0t}^L is the 0.025 posterior quantile for x_{0t}, and x_{0t}^U is the 0.975 posterior quantile for x_{0t}, where x_{0t} is the true value of the calibration target from the second stage of experimentation; then (x_{0t}^L, x_{0t}^U) is a 95% credible interval. The coverage probability (CP) is calculated as such

$$CP = \frac{1}{T}\sum_{t=1}^{T}\psi_t$$

where

$$\psi_t = P[x_{0t}^L < x_{0t} < x_{0t}^U] = \begin{cases} 0 & \text{if } x_{0t} \notin (x_{0t}^L, x_{0t}^U); \\ \\ 1 & \text{if } x_{0t} \in (x_{0t}^L, x_{0t}^U). \end{cases}$$

The average coverage probability ($AvCP$) is calculated by averaging across the number of replications in the simulation study, where

$$AvCP = \frac{1}{100}\sum_{j=1}^{100} CP_j.$$

Another quantity of interest to compare the average interval widths ($AvIW$) for the methods, where the average interval widths (IW) across the simulated time series is calculated as follows:

$$IW = \frac{1}{T}\sum_{t=1}^{T}(x_{0t}^U - x_{0t}^L)$$

with the average interval width across the simulation study given as

$$AvIW = \frac{1}{100}\sum_{j=1}^{100} IW_j,$$

where IW_j is the average interval width for the j^{th} simulation replicate. The performance of the dynamic calibration approaches will be assessed using the average coverage probability ($AvCP$), average interval width ($AvIW$) and average mean square ($AvMSE$).

We consider the performance of the methods under two conditions: interpolation and extrapolation. The interpolation case is of interest to understand how the method performed when the calibrated time series is within the range of the reference values, [20, 100]. The extrapolation case was conducted to examine the methods when x_{0t} falls outside of the range of the calibration references, where $x_{0t} > 100$. While it is not preferable to do extrapolation in the regression case, it is often done in practice in microwave radiometry.

All simulations were carried out on the Compile server running R 3.0.2 (R Development Core Team, 2013) at Virginia Commonwealth University. The Compile server has a Linux OS with 16 CPU cores and 32 GB Ram. Each iteration (the 100 replicates of the simulation) in the study took approximately 15 minutes with a total of 25.63 hours.

2.3.1 Linear Interpolation case

In the following tables, the simulation results for the dynamic and static calibration methods are provided. The results of simulation studies provide insight into the properties of the calibration approaches. The results in Tables 2.1 and 2.2 indicate that all of the estimators do a good job at approximating the true values of x_{0t} when the gain fluctuation g_t is set to 0. Even in this case we see as the variance ratio r increases so does the $AvMSE$ values. All of the methods have an average coverage probability $AvCP$ of 1 or close. The high coverage rate is of no surprise for a stable system. There does not appear to be an advantage by including more reference measurements (i.e 2- or 5-points) in the model when the system is stable in time. The clear difference is the $AvIW$ values for the dynamic methods compared to the static methods. In Tables 2.1 and 2.2 when $r = 10$ and $r = 2$, the interval for the dynamic methods is wider than those of the four static methods but as r increases the interval width of the dynamic methods remain nearly unchanged as the interval widths for the static methods are 4 to 5 times wider.

The simulation results for the stepped gain fluctuations are provided in Tables 2.3 and 2.4. Clearly the presence of the stepped g_t has an effect on the fit of the models. The results in Tables 2.3 show that in nearly all cases, the two dynamic methods M_{D1} and M_{D2} have $AvMSE$ values smaller than the two static Bayesian approaches. The $AvMSE$ values for the dynamic methods are reasonably lower for $r = 200$. When $r = 1000$, notice the dynamic models M_{D1} and M_{D2} have smaller average mean square errors smaller than the static method M_{F2}. The average coverage probability $AvCP$ is comparable for all of the

methods and number of references. The dynamic methods consistently have shorter interval widths. The widths of the 95% credible intervals for M_{D1} and M_{D2} is not affected by the increases in r.

The results provided in Tables 2.5 and 2.6 summarize the performance of the methods

Table 2.1: *Comparison of calibration approaches when interpolating to estimate x_{0t} without gain fluctuations.*

Constant $g_t = 0$

		$r = 10$			$r = 100$			$r = 1000$		
Ref.	Model	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW
2	M_{D1}	0.0008	0.995	2.519	0.0035	0.983	2.523	0.0307	0.939	2.517
	M_{D2}	0.0012	1.000	3.782	0.0038	1.000	3.782	0.0308	1.000	3.782
	M_{F1}	0.0001	1.000	1.224	0.0012	1.000	3.868	0.0123	1.000	12.229
	M_{F2}	0.0001	1.000	1.223	0.0016	1.000	3.863	0.0335	1.000	12.168
	M_{B1}	0.0002	0.997	1.182	0.0022	1.000	3.866	0.0386	1.000	12.177
	M_{B2}	0.0014	1.000	1.458	0.0139	1.000	4.606	0.1391	1.000	14.565
5	M_{D1}	0.0008	0.995	2.496	0.0035	0.983	2.509	0.0307	0.941	2.514
	M_{D2}	0.0013	1.000	3.983	0.0039	1.000	3.983	0.0307	1.000	3.983
	M_{F1}	0.0001	1.000	1.223	0.0012	1.000	3.865	0.0123	1.000	12.220
	M_{F2}	0.0001	1.000	1.222	0.0022	1.000	3.860	0.0813	1.000	12.113
	M_{B1}	0.0002	1.000	1.223	0.0023	1.000	3.861	0.0792	1.000	12.116
	M_{B2}	0.0014	1.000	1.457	0.0139	1.000	4.604	0.1069	1.000	10.748

when the gain fluctuation is sinusoidal noise. The results for r values of 10, 100, and 1000 are given in Table 2.5 with $r = 2,\ 20$ and 200 given in Table 2.6. When g_t is sinusoidal, the $AvMSE$ values for the dynamic methods are consistently larger than any of the static methods. For all of the chosen r values, the $AvCP$ is considerably lower than the opposing methods. The dynamic methods still have average interval widths extremely shorter than any of the static methods. The $AvIW$ is constant across the variance ratios.

The simulation study shows that methods M_{D1} and M_{D2} do a good job at estimating calibrated values that are interior to the range of reference measurements. Both methods display high coverage probabilities in the presence of drifting parameters. For the three possible gain fluctuations, the interval widths for the dynamic methods were consistently shorter than the static calibration approaches. When fitting data where there is a definite linear relationship, the dynamic methods are invariant to the number of reference measurements. When using the proposed methods in this paper, not much will be gained by using more than 2 reference measurements. Overall, when interpolating to estimate x_{0t}, the dynamic methods outperform the static Bayesian approaches across the different variance ratios. In the following section the performance of the dynamic methods are assessed when the calibrated values fall outside of the range of reference measurements.

Table 2.2: *Comparison of calibration approaches when interpolating to estimate x_{0t} without gain fluctuations.*

Constant $g_t = 0$

		$r = 2$			$r = 20$			$r = 200$		
Ref.	Model	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW
2	M_{D1}	0.0012	0.992	2.519	0.0041	0.981	2.520	0.0323	0.939	2.528
	M_{D2}	0.0015	1.000	3.782	0.0044	1.000	3.782	0.0325	1.000	3.782
	M_{F1}	0.0001	1.000	1.230	0.0010	1.000	3.871	0.0114	1.000	12.231
	M_{F2}	0.0001	1.000	1.229	0.0012	1.000	3.866	0.0314	1.000	12.170
	M_{B1}	0.0001	1.000	1.230	0.0019	1.000	3.869	0.0371	1.000	12.179
	M_{B2}	0.0190	1.000	1.155	0.0243	1.000	3.767	0.1381	1.000	14.567
5	M_{D1}	0.0011	0.992	2.508	0.0041	0.981	2.510	0.032	0.939	2.514
	M_{D2}	0.0017	1.000	3.983	0.0045	1.000	3.983	0.032	1.000	3.983
	M_{F1}	0.0001	1.000	1.228	0.0010	1.000	3.868	0.011	1.000	12.222
	M_{F2}	0.0001	1.000	1.227	0.0019	1.000	3.863	0.081	1.000	12.114
	M_{B1}	0.0001	1.000	1.227	0.0021	1.000	3.864	0.082	1.000	12.118
	M_{B2}	0.0013	1.000	1.462	0.0137	1.000	4.607	0.138	1.000	14.560

Table 2.3: *Comparison of calibration approaches when interpolating to estimate x_{0t} with stepped gain fluctuations.*

Stepped $g_t = a_i$

		$r = 10$			$r = 100$			$r = 1000$		
Ref.	Model	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW
2	M_{D1}	0.0191	0.961	2.509	0.0198	0.953	2.506	0.0406	0.926	2.543
	M_{D2}	0.0196	1.000	3.782	0.0201	1.000	3.782	0.0408	1.000	3.783
	M_{F1}	0.0001	1.000	9.094	0.0004	1.000	9.813	0.0094	1.000	15.209
	M_{F2}	0.0046	1.000	9.065	0.0073	1.000	9.779	0.0528	1.000	15.098
	M_{B1}	0.0859	1.000	9.072	0.0838	1.000	9.786	0.1866	1.000	15.109
	M_{B2}	0.1399	1.000	10.830	0.0823	1.000	11.687	0.1836	1.000	18.115
5	M_{D1}	0.0191	0.961	2.510	0.0197	0.954	2.511	0.0405	0.924	2.516
	M_{D2}	0.0196	1.000	3.983	0.0201	1.000	3.983	0.0405	1.000	3.983
	M_{F1}	0.0001	1.000	9.087	0.0004	1.000	9.806	0.0094	1.000	15.199
	M_{F2}	0.0184	1.000	9.041	0.0267	1.000	9.749	0.1620	1.000	14.995
	M_{B1}	0.0199	1.000	9.044	0.0267	1.000	9.752	0.1559	1.000	14.999
	M_{B2}	0.0706	1.000	10.826	0.0618	1.000	8.625	0.1742	1.000	15.091

2.3.2 Extrapolation case

At this point in the paper we examine the calibration approaches when the calibrated values are outside of the reference measurements. The range of the measurement references is

Table 2.4: *Comparison of calibration approaches when interpolating to estimate x_{0t} with stepped gain fluctuations.*

Stepped $g_t = a_i$

		$r = 2$			$r = 20$			$r = 200$		
Ref.	Model	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW
2	M_{D1}	0.0209	0.957	2.520	0.0219	0.950	2.522	0.0436	0.921	2.526
	M_{D2}	0.0214	1.000	3.782	0.0222	1.000	3.782	0.0438	1.000	3.782
	M_{F1}	0.0001	1.000	9.103	0.0003	1.000	9.822	0.0086	1.000	15.216
	M_{F2}	0.0047	1.000	9.075	0.0073	1.000	9.788	0.0511	1.000	15.105
	M_{B1}	0.0084	1.000	9.081	0.0115	1.000	9.795	0.0601	1.000	15.116
	M_{B2}	0.0709	1.000	10.842	0.0826	1.000	11.698	0.2054	1.000	18.122
5	M_{D1}	0.0209	0.957	2.509	0.0218	0.949	2.511	0.0436	0.920	2.516
	M_{D2}	0.0214	1.000	3.983	0.0221	1.000	3.983	0.0435	1.000	3.983
	M_{F1}	0.0001	1.000	9.096	0.0003	1.000	9.815	0.0086	1.000	15.205
	M_{F2}	0.0185	1.000	9.050	0.0267	1.000	9.758	0.1616	1.000	15.002
	M_{B1}	0.0199	1.000	9.053	0.0281	1.000	9.761	0.1641	1.000	15.006
	M_{B2}	0.0708	1.000	10.836	0.0825	1.000	11.693	0.2052	1.000	18.114

Table 2.5: *Comparison of calibration approaches when interpolating to estimate x_{0t} with sinusoidal gain fluctuations.*

Sinusoidal $g_t = 0.1\sin(0.025t)$

		$r = 10$			$r = 100$			$r = 1000$		
Ref.	Model	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW
2	M_{D1}	4.4088	0.628	2.657	4.4794	0.629	2.648	4.7214	0.638	2.681
	M_{D2}	4.4002	0.829	3.783	4.4708	0.825	3.783	4.7123	0.810	3.783
	M_{F1}	0.0001	1.000	21.980	0.0012	1.000	22.307	0.0123	1.000	25.206
	M_{F2}	0.1541	1.000	21.665	0.1670	1.000	21.978	0.2943	1.000	24.738
	M_{B1}	0.1689	0.975	20.933	0.1868	1.000	21.994	0.3174	1.000	24.757
	M_{B2}	0.4127	1.000	26.178	0.4258	1.000	26.567	0.5531	1.000	30.020
5	M_{D1}	4.4087	0.628	2.646	4.4793	0.630	2.648	4.7214	0.635	2.653
	M_{D2}	4.3906	0.845	3.984	4.4609	0.839	3.984	4.7023	0.824	3.984
	M_{F1}	0.0001	1.000	21.964	0.0012	1.000	22.291	0.0123	1.000	25.188
	M_{F2}	0.5810	1.000	21.371	0.6218	1.000	21.671	1.0152	1.000	24.306
	M_{B1}	0.5956	1.000	21.377	0.5909	1.000	21.678	0.9628	1.000	24.314
	M_{B2}	0.4123	1.000	26.166	0.3087	1.000	18.973	0.4658	1.000	25.009

from 20 to 100. The true x_{0t} behaved as a random walk bounded between 100 and 110. We assessed the performance of the dynamic methods under three possible gain fluctuation patterns. First, the simulation study is conducted without the presence of additional gain

Table 2.6: *Comparison of calibration approaches when interpolating to estimate x_{0t} with sinusoidal gain fluctuations.*

Sinusoidal $g_t = 0.1\sin(0.025t)$

		$r = 2$			$r = 20$			$r = 200$		
Ref.	Model	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW
2	M_{D1}	4.4504	0.625	2.658	4.5213	0.628	2.660	4.7643	0.6331	2.665
	M_{D2}	4.4419	0.828	3.783	4.5127	0.823	3.783	4.7553	0.8083	3.783
	M_{F1}	0.0001	1.000	21.968	0.0010	1.000	22.295	0.0114	1.000	25.196
	M_{F2}	0.1538	1.000	21.653	0.1672	1.000	21.966	0.2896	1.000	24.729
	M_{B1}	0.1732	1.000	21.669	0.1867	1.000	21.982	0.3127	1.000	24.748
	M_{B2}	0.4122	1.000	26.164	0.4253	1.000	26.553	0.5518	1.000	30.008
5	M_{D1}	4.4504	0.625	2.647	4.5213	0.627	2.648	4.7643	0.633	2.654
	M_{D2}	4.4322	0.844	3.984	4.5029	0.838	3.984	4.7451	0.823	3.984
	M_{F1}	0.0001	1.000	21.952	0.0010	1.000	22.278	0.0114	1.000	25.178
	M_{F2}	0.5799	1.000	21.359	0.6206	1.000	21.660	1.0133	1.000	24.297
	M_{B1}	0.5851	1.000	21.366	0.6256	1.000	21.667	1.0183	1.000	24.304
	M_{B2}	0.4119	1.000	26.152	0.4247	1.000	26.541	0.5513	1.000	29.995

fluctuation (i.e. $g_t = 0$); second, the gain g_t is a stepped pattern influencing the time-varying slope β_{1t} over time; lastly, a sinusoidal g_t is added to β_{1t}. Just as the previous results, the methods are assessed by the average mean square error ($AvMSE$), average coverage probability ($AvCP$), and the average interval width ($AvIW$) under different signal-noise-ratios.

The results are provided in Tables 2.7 and 2.8 for the statistical calibration methods without gain fluctuations. The performance of the proposed method is stable across the variance ratios. A point of interest is the reported $AvIW$ values for methods M_{D1} and M_{D2}. We see for $r = 10$ and $r = 2$ that the $AvIW$ is 3 to 5 times wider than those for the static approaches. When $r = 100$ and $r = 20$ the interval width for all competing methods are relatively close. The dynamic approaches outperform the static methods in noisy conditions such as $r = 1000$ and $r = 200$. The interval widths for the dynamic methods are considerably shorter than the those for the static methods. The simulation results reveal that when the data is characteristic of having a large variance ratio, the dynamic methods, M_{D1} and M_{D2}, will outperform static Bayesian approaches and the inverse approach.

Next, we impose a stepped gain fluctuation g_t to the data generated and wanted to evaluate the behavior of the calibration methods. The results for the stepped case are given in Tables 2.9 and 2.10. We see by the $AvMSE$ values in both tables that the dynamic methods perform better than most static methods. If the calibrated values by chance drift outside of the reference range the dynamic methods will do a good job at capturing them with certainty while having a narrower credible interval than confidence intervals of the static methods. The dynamic approaches outperform all of the static method in terms of $AvIW$.

Table 2.7: *Comparison of calibration approaches when extrapolating to estimate x_{0t} without gain fluctuations.*

Constant $g_t = 0$

		$r = 10$			$r = 100$			$r = 1000$		
Ref.	Model	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW
2	M_{D1}	0.0018	1.000	5.255	0.0043	1.000	5.255	0.0309	1.000	5.255
	M_{D2}	0.0016	1.000	3.910	0.0042	1.000	3.910	0.0311	1.000	3.910
	M_{F1}	0.0001	1.000	1.224	0.0012	1.000	3.869	0.0123	1.000	12.234
	M_{F2}	0.0001	1.000	1.223	0.0001	1.000	3.863	0.1019	1.000	12.168
	M_{B1}	0.0001	1.000	1.225	0.0008	1.000	3.867	0.1115	1.000	12.181
	M_{B2}	0.0014	1.000	1.458	0.0139	1.000	4.606	0.1391	1.000	14.565
5	M_{D1}	0.0019	1.000	5.233	0.0043	1.000	5.233	0.0309	1.000	5.233
	M_{D2}	0.0029	1.000	4.106	0.0054	1.000	4.106	0.0323	1.000	4.106
	M_{F1}	0.0001	1.000	1.223	0.0012	1.000	3.866	0.0123	1.000	12.224
	M_{F2}	0.0001	1.000	1.222	0.0027	1.000	3.860	0.5502	1.000	12.113
	M_{B1}	0.0001	1.000	1.223	0.0030	1.000	3.862	0.5566	1.000	12.120
	M_{B2}	0.0014	1.000	1.457	0.0139	1.000	4.604	0.1389	1.000	14.558

Table 2.8: *Comparison of calibration approaches when extrapolating to estimate x_{0t} without gain fluctuations.*

Constant $g_t = 0$

		$r = 2$			$r = 20$			$r = 200$		
Ref.	Model	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW
2	M_{D1}	0.0031	1.000	5.253	0.0060	1.000	5.253	0.0340	1.000	5.253
	M_{D2}	0.0034	1.000	3.910	0.0064	1.000	3.910	0.0347	1.000	3.910
	M_{F1}	0.0001	1.000	1.230	0.0008	1.000	3.872	0.0107	1.000	12.236
	M_{F2}	0.0001	1.000	1.229	0.0003	1.000	3.866	0.1068	1.000	12.170
	M_{B1}	0.0001	1.000	1.230	0.0010	1.000	3.871	0.1164	1.000	12.183
	M_{B2}	0.0013	1.000	1.465	0.0135	1.000	4.610	0.1376	1.000	14.567
5	M_{D1}	0.0032	1.000	5.231	0.0060	1.000	5.231	0.0340	1.000	5.232
	M_{D2}	0.0053	1.000	4.106	0.0083	1.000	4.106	0.0365	1.000	4.106
	M_{F1}	0.0001	1.000	1.228	0.0008	1.000	3.869	0.0107	1.000	12.226
	M_{F2}	0.0001	1.000	1.227	0.0035	1.000	3.863	0.5616	1.000	12.114
	M_{B1}	0.0001	1.000	1.228	0.0039	1.000	3.865	0.5680	1.000	12.121
	M_{B2}	0.0013	1.000	1.462	0.0135	1.000	4.607	0.1375	1.000	14.560

These results of the simulation study do not change much across the number of references used. Once again, when the relationship is assumed to be linear there is no benefit to adding more references.

Table 2.9: *Comparison of calibration approaches when extrapolating to estimate x_{0t} with stepped gain fluctuations.*

Stepped $g_t = a_i$

		$r = 10$			$r = 100$			$r = 1000$		
Ref.	Model	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW
2	M_{D1}	0.0206	1.000	5.247	0.0210	1.000	5.247	0.0412	1.000	5.247
	M_{D2}	0.0225	1.000	3.910	0.0230	1.000	3.910	0.0435	1.000	3.910
	M_{F1}	0.0001	1.000	9.097	0.0004	1.000	9.817	0.0094	1.000	15.215
	M_{F2}	0.0581	1.000	9.065	0.0656	1.000	9.779	0.3191	1.000	15.098
	M_{B1}	0.0634	1.000	9.075	0.0718	1.000	9.789	0.3361	1.000	15.115
	M_{B2}	0.0707	1.000	10.830	0.0826	1.000	11.687	0.2060	1.000	18.115
5	M_{D1}	0.0209	1.000	5.226	0.0213	1.000	5.226	0.0412	1.000	5.226
	M_{D2}	0.0268	1.000	4.106	0.0273	1.00	4.106	0.0483	1.000	4.106
	M_{F1}	0.0001	1.000	9.090	0.0004	1.000	9.809	0.0094	1.000	15.203
	M_{F2}	0.2274	1.000	9.041	0.2812	1.000	9.749	1.4628	1.000	14.995
	M_{B1}	0.2307	1.000	9.047	0.2851	1.000	9.755	1.4744	1.000	15.004
	M_{B2}	0.0706	1.000	10.826	0.0825	1.000	11.682	0.2058	1.000	18.106

Table 2.10: *Comparison of calibration approaches when extrapolating to estimate x_{0t} with stepped gain fluctuations.*

Stepped $g_t = a_i$

		$r = 2$			$r = 20$			$r = 200$		
Ref.	Model	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW
2	M_{D1}	0.0242	1.000	5.245	0.0250	1.000	5.245	0.0466	1.000	5.245
	M_{D2}	0.0266	1.000	3.910	0.0275	1.000	3.910	0.0494	1.000	3.910
	M_{F1}	0.0001	1.000	9.106	0.0002	1.000	9.826	0.0080	1.000	15.222
	M_{F2}	0.0620	1.000	9.075	0.0698	1.000	9.788	0.3284	1.000	15.105
	M_{B1}	0.0674	1.000	9.085	0.0760	1.000	9.799	0.3447	1.000	15.121
	M_{B2}	0.0710	1.000	10.842	0.0825	1.000	11.698	0.2048	1.000	18.122
5	M_{D1}	0.0245	1.000	5.224	0.0254	1.000	5.224	0.0427	1.000	5.226
	M_{D2}	0.0315	1.000	4.106	0.0324	1.000	4.106	0.0485	1.000	4.106
	M_{F1}	0.0001	1.000	9.099	0.0002	1.000	9.818	0.0089	1.000	14.255
	M_{F2}	0.2354	1.000	9.050	0.2902	1.000	9.758	1.1896	1.000	14.078
	M_{B1}	0.2388	1.000	9.056	0.2941	1.000	9.764	1.1995	1.000	14.086
	M_{B2}	0.0709	1.000	10.836	0.0824	1.000	11.693	0.1831	1.000	16.977

Lastly, the study is conducted with a sinusoidal gain fluctuation while extrapolating to estimate x_{0t}. The results for the sinusoidal case are given in Tables 2.11 and 2.12. The dynamic methods M_{D1} and M_{D2} exhibit the same behavior as before in Tables 2.5 and 2.6

Table 2.11: *Comparison of calibration approaches when extrapolating to estimate x_{0t} with sinusoidal gain fluctuations.*

Sinusoidal $g_t = 0.1\sin(0.025t)$

		$r = 10$			$r = 100$			$r = 1000$		
Ref.	Model	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW
2	M_{D1}	4.4096	0.873	5.127	4.4800	0.872	5.127	4.7214	0.866	5.127
	M_{D2}	4.4410	0.833	3.904	4.5114	0.825	3.904	4.7530	0.813	3.904
	M_{F1}	0.0001	1.000	21.988	0.0012	1.000	22.315	0.0123	1.000	25.216
	M_{F2}	1.8193	1.000	21.665	1.8636	1.000	21.978	2.7760	1.000	24.739
	M_{B1}	1.8602	1.000	21.688	1.9056	1.000	22.002	2.8312	1.000	24.766
	M_{B2}	0.4127	1.000	26.178	0.4258	1.000	26.567	0.5531	1.000	30.020
5	M_{D1}	4.4105	0.872	5.106	4.4808	0.872	5.106	4.7222	0.866	5.107
	M_{D2}	4.4889	0.842	4.100	4.5593	0.835	4.101	4.8007	0.822	4.100
	M_{F1}	0.0001	1.000	21.971	0.0012	1.000	22.297	0.0123	1.000	25.195
	M_{F2}	6.9539	1.000	21.371	7.2327	1.000	21.671	11.0337	1.000	24.306
	M_{B1}	6.9852	1.000	21.383	7.2650	1.000	21.684	11.0772	1.000	24.320
	M_{B2}	0.4123	1.000	26.166	0.4254	1.000	26.555	0.5526	1.000	30.007

Table 2.12: *Comparison of calibration approaches when extrapolating to estimate x_{0t} with sinusoidal gain fluctuations.*

Sinusoidal $g_t = 0.1\sin(0.025t)$

		$r = 2$			$r = 20$			$r = 200$		
Ref.	Model	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW	AvMSE	AvCP	AvIW
2	M_{D1}	4.4491	0.872	5.125	4.5199	0.871	5.125	4.7626	0.866	5.126
	M_{D2}	4.4809	0.828	3.904	4.5518	0.821	3.904	4.7948	0.807	3.904
	M_{F1}	0.0001	1.000	21.976	0.0008	1.000	22.303	0.0107	1.000	25.205
	M_{F2}	1.8350	1.000	21.653	1.8796	1.000	21.966	2.7956	1.000	24.729
	M_{B1}	1.8759	1.000	21.676	1.9216	1.000	21.990	2.8508	1.000	24.756
	M_{B2}	0.4123	1.000	26.164	0.4250	1.000	26.553	0.5511	1.000	30.008
5	M_{D1}	4.4497	0.872	5.10	4.5205	0.871	5.105	4.7633	0.865	5.105
	M_{D2}	4.5292	0.836	4.100	4.6000	0.832	4.100	4.842	0.814	4.100
	M_{F1}	0.0001	1.000	21.958	0.0008	1.000	22.285	0.0107	1.000	25.185
	M_{F2}	6.9764	1.000	21.359	7.2560	1.000	21.660	11.0629	1.000	24.297
	M_{B1}	7.0077	1.000	21.372	7.2882	1.000	21.673	11.1064	1.000	24.311
	M_{B2}	0.4119	1.000	26.152	0.4246	1.000	26.541	0.5507	1.000	29.995

with $AvMSE$ values ranging for 4.4 to 4.8. Even though the average mean square errors are larger than those of the static methods when using a 2-reference model, the two dynamic methods outperform the static methods M_{F2} and M_{B1}, which are based on the inverse

approach. The dynamic models have average coverage probabilities smaller than the static model across all of the variance ratios. We cannot fail to point out that once again the $AvIW$ are 4 to 6 times shorter than the average widths for the static models.

2.4 Application to Microwave Radiometry

In this example, we apply the dynamic calibration approaches to the calibration of a microwave radiometer for an earth observing satellite. Engineers and scientist commonly use microwave radiometers to measure the electromagnetic radiation emitted by some source or a particular surface such as ice or land surface. Radiometers are very sensitive instruments that are capable of measuring extremely low levels of radiation. The transmission source of the radiant power is the target of the radiometers antenna. When the region of interest, such as terrain, is observed by a microwave radiometer, the radiation received by the antenna is partly due to self-emission by the area of interest and partly due to the reflected radiation originating from the surroundings (Ulaby *et al.* 1981), such as cosmic background radiation, ocean surface, or a heated surface used for the purpose of calibration.

A basic diagram of a radiometer is shown in Figure 2.2 where the radiant power with equivalent brightness temperature (the intensity of the radiation emitted by the scene under observation) T_A enters the radiometer receiver and is converted to the output signal $v(t)$. The schematic features the common components of most microwave radiometers. As the

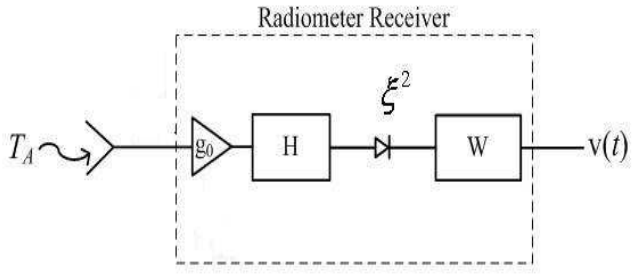


Figure 2.2: *Schematic of Simple Radiometer*

radiometer captures a signal (i.e. Brightness Temperature T_A), it couples the signal into a transmission line which then carries the signal to and from the various elements of the circuit. In Figure 2.2, a signal T_A is introduced directly into the antenna, then it is mixed, amplified and filtered to produce the output signal $v(t)$. This filtering and amplification of the signal is carried out through the following components of the radiometer: an amplifier (g_0); pre-detection filter (H); a square law detector (ξ^2); and a post-detection filter (W). The output of the radiometer is denoted as $v(t)$. See Ulaby *et al.* (1981) for a detailed discussion.

Racette and Lang (2005) state that at the core of every radiometer measurement is a cal-

ibrated receiver. Calibration is required due to the fact that the current electronic hardware is unable to maintain a stable input/output relationship. For space observing instruments, stable calibration without any drifts is a key to detect proper trends of climate (Imaoka *et al.* 2010). Problems such as amplifier gain instability and exterior temperature variations of critical components may cause this relationship to drift over time (Bremer 1979). During the calibration process, the radiometer receiver measures the voltage output power $v(t)$, and its corresponding input temperature of a known reference. Two or more known reference temperatures are needed for calibration of a radiometer. Ulaby *et al.* (1981), and Racette and Lang (2005) state that the relationship between the output, $v(t)$, and the input, T_A, is approximately linear, and can be expressed as

$$\hat{T}_A = \beta_0 + \beta_1 v(t)$$

where, $\hat{T}_A$ is the estimated value of the brightness temperature, $v(t)$ is the observed output voltage. Using this relationship, the output value, $v(t)$, is used to derive an estimate for the input, T_A (Racette and Lang, 2005).

Traditional calibration methods use measurements taken from known calibration ref-

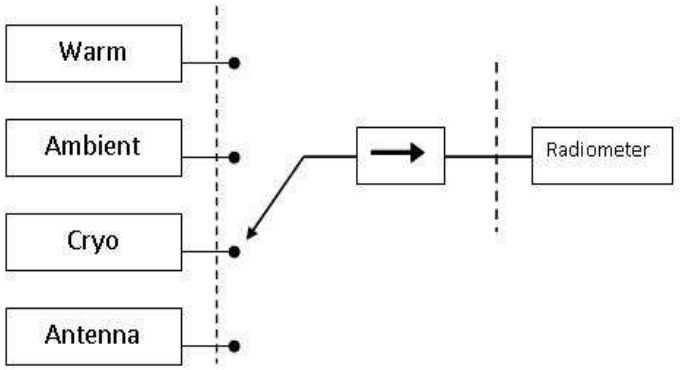


Figure 2.3: *Known Reference Temperature Collection*

erences, for example see Figure 2.3. Due to possible cost constraints it is common to use between two and five references. The reference temperatures are converted to their equivalent power measurement prior to the calibration algorithm. The radiometer outputs are observed when the radiometer measures the reference temperatures, giving an ordered calibration pair (T_i, v_i). The v_i values are observed from the process of the electronics within the radiometer (see Figure 2.2) (Ulaby *et al.* 1981; Racette and Lang 2005). Through the process of calibration, the unknown brightness temperature T_j is estimated by plugging its observed output v_j into either Equation (2.3) or Equation (2.4).

It is of interest to develop a calibration approach that can detect gain abnormalities, and/or correct for slow drifts that affect the quality of the instrument measurements. To demonstrate the dynamic approach in terms of application appeal, the two dynamic methods were used to characterize a calibration target over time for a microwave radiometer. The

data used for this example was collected during a calibration experiment that was conducted on the Millimeter-wave Imaging Radiometer (MIR) (Racette et al. 1995). The purpose of the experiment was to validate predictions of radiometer calibration.

The MIR was built with two internal blackbody references which will be used to observe

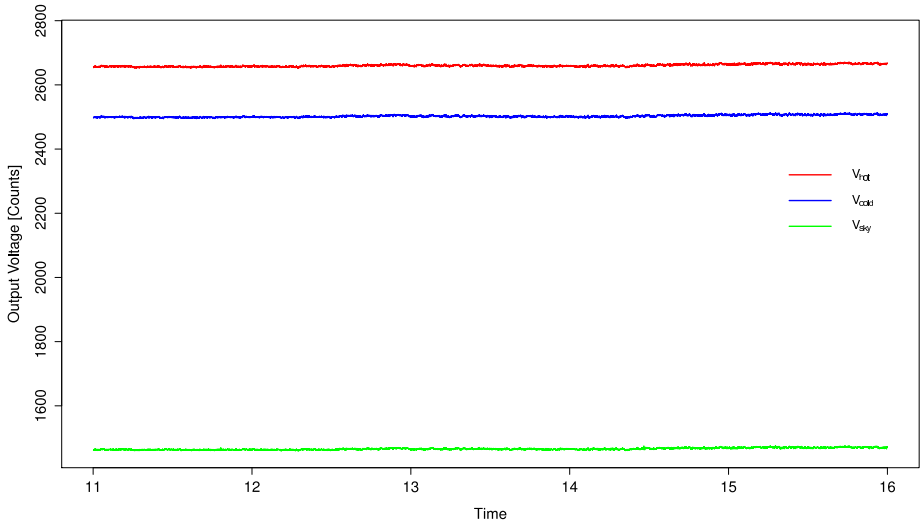


Figure 2.4: *Time series of MIR output voltage measurement data V_{cold}, V_{hot}, and V_{sky}.*

a third stable temperature reference for an extended period of time. The third reference was a custom designed cryogenically cooled reference. Racette (2005) conducted the MIR experiment under two scenarios: the first experiment denoted as $T295$ examined the calibration predictions when the unknown target is interior (i.e. interpolation) to the reference measurements; the second set of measurements (denoted as $T80$) were taken when the unknown temperature to be estimated is outside (i.e. extrapolation) of the range of calibration references.

For demonstration purposes we will only consider the $T80$ experiment; for details of the $T295$ experiment see Racette (2005). For the $T80$ run of the experiment, the reference temperatures are as follows:

1. $T_{cold} \sim 293.69K$
2. $T_{hot} \sim 325.59K$

with the unknown target temperature that must be estimated denoted as T_{sky}. Each temperature measure has a corresponding observed time series of output measurements; V_{cold}, V_{hot}, and V_{sky} (see Figure 2.4). Therefore in this example we only consider a 2-point calibration set-up as we use T_{cold} and T_{hot} as the known reference standards and use V_{sky} to derive

estimates of T_{sky} for the first 1000 time periods.

The results of the dynamic approaches: M_{D1} and M_{D2}, will be compared to the "inverse" calibration method (Krutchkoff 1967) implemented by Racette (2005). The method considered by Racette (2005) will be denoted as M_{1u}. As in practice, rarely does one know the value of the true temperature to be estimated so the aim of this example is to assess the contribution of the calibration approach to the variability in the measurement estimate. The Racette (2005) analysis did not consider biases that may exist in calibration, and continuing in the same spirit, the existence of biases will not be considered in the analysis. We will apply the M_{1u}, M_{D1}, and M_{D2} approaches to the data to estimate the temperature T_{sky}; the standard deviation of the estimated time series $\hat{\sigma}_{T_{sky}}$ is used as a measure of uncertainty, including the contribution of the calibration algorithm.

Figure 2.5 shows the time series of the temperature estimates for T_{sky} using Krutchkoff's

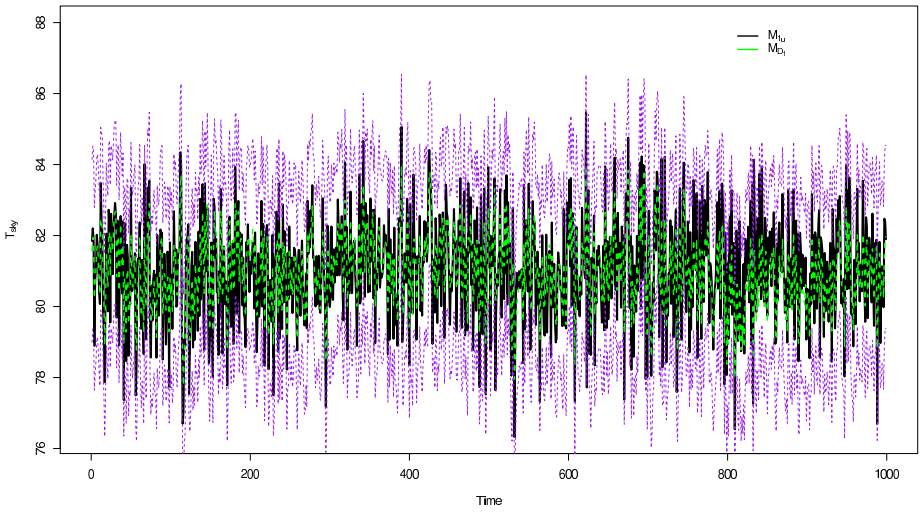


Figure 2.5: *Time series of calibrated temperature for MIR T80 experiment. The "inverse" calibration approach M_{1u} vs the dynamic approach M_{D1} with the 95% credible intervals.*

(1967) "inverse" approach M_{1u} and the dynamic approach M_{D1}. The standard deviations for the M_{1u} and M_{D1} approaches are $\hat{\sigma}_{T_{sky}}(M_{1u}) = 1.482K$ and $\hat{\sigma}_{T_{sky}}(M_{D1}) = 0.998K$, respectfully. We see the dynamic model M_{D1} improves the estimation process over the static model M_{1u} by observing the corresponding standard deviation values. The dynamic model decreased the measurement uncertainty by roughly 33%. In Figure 2.6 the time series of the temperature estimates for T_{sky} using the "inverse" approach M_{1u} and the dynamic approach M_{D2} is given. The standard deviations for the M_{1u} and M_{D2} approaches are $\hat{\sigma}_{T_{sky}}(M_{1u}) = 1.482K$ and $\hat{\sigma}_{T_{sky}}(M_{D2}) = 0.974K$. Again, the dynamic approach outperforms the static model M_{1u}. In this case, dynamic model M_{D2} decreased the measurement uncertainty by roughly 34%.

We assess the normality assumption that the error vector $\boldsymbol{\epsilon}_t$, are normally distributed random variables with mean equal 0 and a constant variance σ^2. The error terms are calcu-

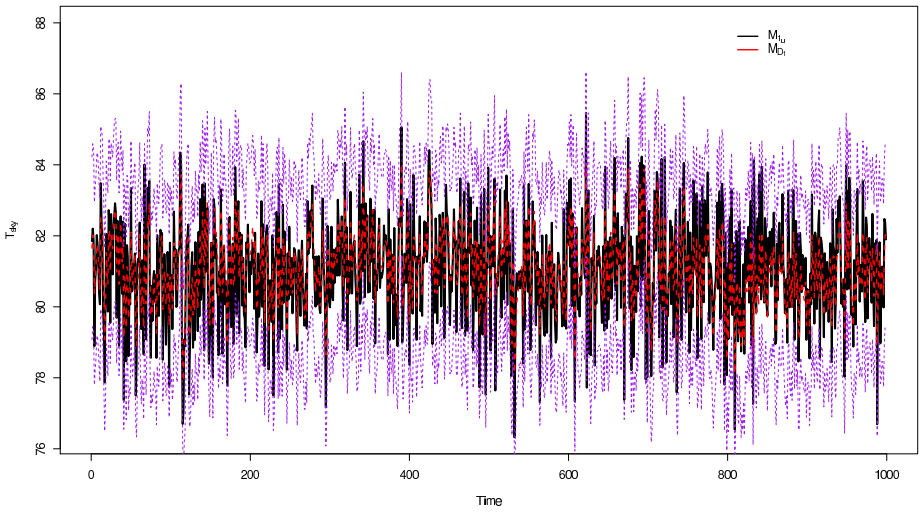


Figure 2.6: *Time series of calibrated temperature for MIR T80 experiment. The "inverse" calibration approach M_{1u} vs the dynamic approach M_{D2} with the 95% credible intervals.*

lated when deriving the posterior distribution at time t for $\boldsymbol{\theta}_t|D_t$ and is denoted as $\mathbf{e}_t$. In Figure 2.7 we provide the Normal Q-Q Plots for $\mathbf{e}_t = [\epsilon_{t_{cold}} \ \epsilon_{t_{hot}}]'$ where $\epsilon_{t_{cold}}$ and $\epsilon_{t_{hot}}$ are the univariate error vectors derived from the DLM framework. We observe that the points all lie quite close to the line; close enough to say these data come from a normal distribution. There isn't much of a random deviation in either Q-Q plots about the line except for near the tails; this does not disqualify these data from being normal. In Figure 2.8, the histograms of $\epsilon_{t_{cold}}$ and $\epsilon_{t_{hot}}$ are given. We see that the residuals, $\epsilon_{t_{cold}}$ and $\epsilon_{t_{hot}}$, produced by the dynamic calibration process are approximately normally distributed. West and Harrison (1997) state that under the assumptions of the dynamic linear model, each error ϵ_t has the predictive distribution

$$(\epsilon_t|D_{t-1}) \sim T_{n_{t-1}}[0, Q_t].$$

They further conclude that when n_{t-1} is large, these distributions are essentially normal.

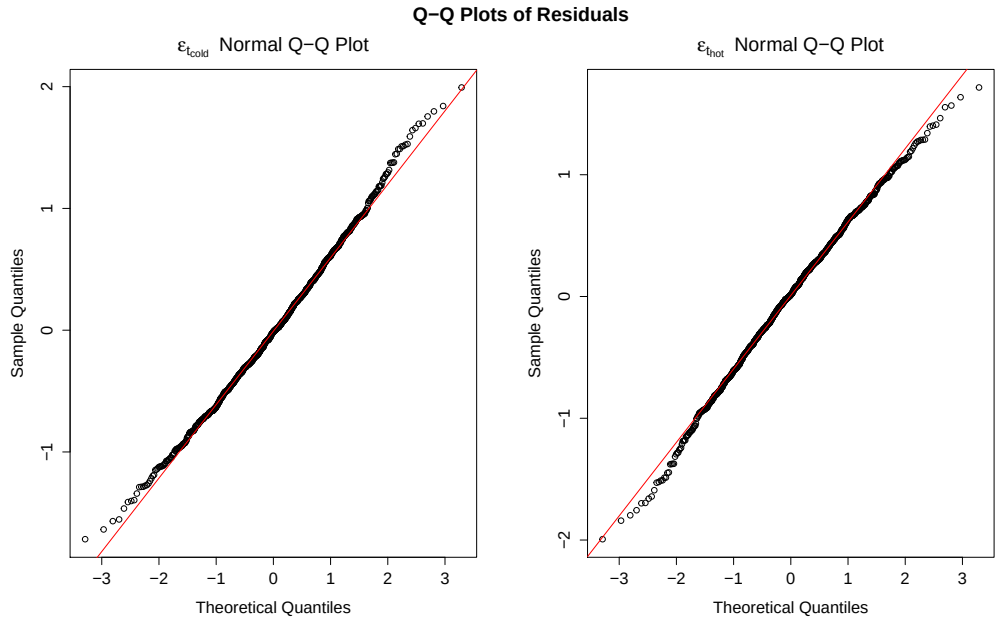


Figure 2.7: *Normal Q-Q Plots of* $\epsilon_{t_{cold}}$ *and* $\epsilon_{t_{hot}}$

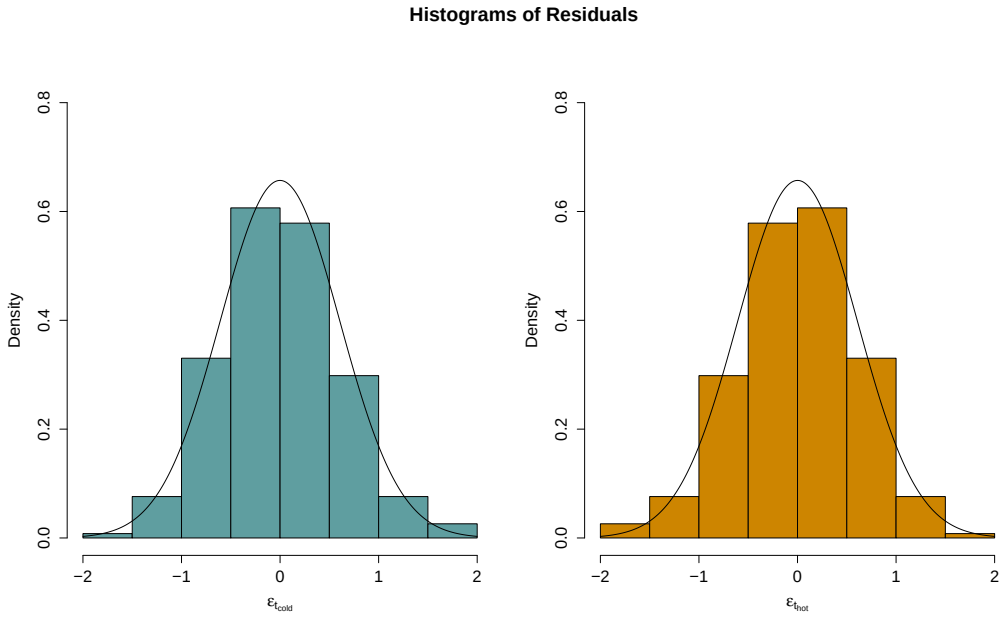


Figure 2.8: *Histograms of* $\epsilon_{t_{cold}}$ *and* $\epsilon_{t_{hot}}$

2.5 Discussion

Two new novel approaches to the statistical calibration problem have been presented in this paper. In was shown by the simulation results that the use of the dynamic approach has its benefits over the static methods. If the linear relationship in the first stage of calibration is known to be stable, then the traditional methods should be used. The dynamic methods showed promise in the cases when the variance ratio was high. There is also a computation expense to implementing the dynamic methods compared to the static methods, but in the sense of electronics these methods allow for near real time calibration and monitoring.

It is worth noting that the dynamic method shows possible deficiencies when the gain fluctuation is sinusoidal, referring to results in Table 2.5. In Figure 2.9, it is evident the largest source of the error is in the beginning of estimation process, roughly from $t = 1$ to $t = 200$. The MSE values for the dynamic approaches; M_{D1} and M_{D2} were 4.41 and 4.40, respectively, which was vastly different than those reported for the static methods. This problem can be addressed by extending the burn-in period. This brief learning period ("burn-in period") is needed because the prior distribution may start the estimation inaccurately.

We increased the burn-in period to 200 which allowed the algorithm more time to learn

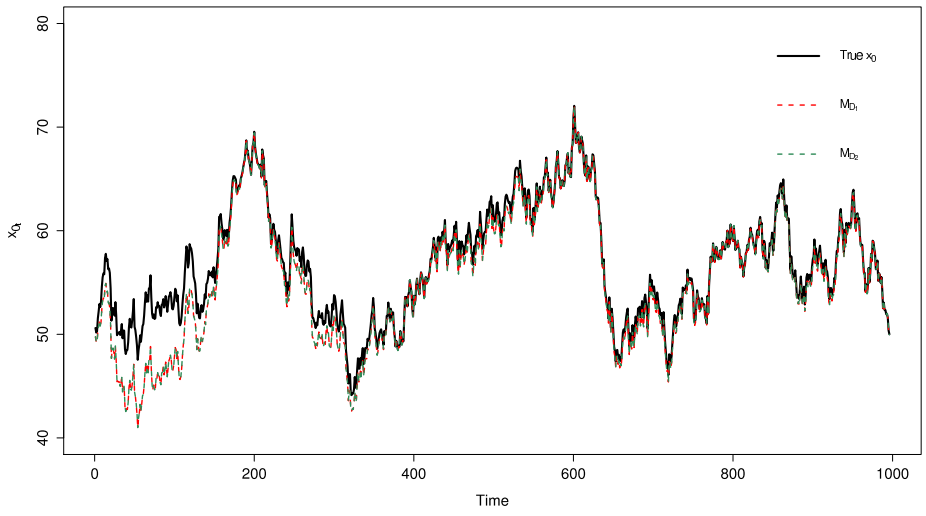


Figure 2.9: *Time series of M_{D1} and M_{D2} estimates in the interpolation case with gain fluctuations $g_t = 0.1sin(0.025t)$ (burn-in = 0).*

and hence results in a lower MSE value. In Figure 2.10 we see that the estimates fit better to the true values of x_{0t}. The MSE decreased from 4.41 to 0.64 for M_{D1} and 0.63 for M_{D2}. The increased burn-in period improves the coverage probability but the interval width isn't noticeably affected. The coverage probability increased from 0.628 to 0.722 for M_{D1} and from 0.829 to 0.964 for M_{D2}.

For completeness we consider the behavior of the method if β_{1t} crosses zero. It is

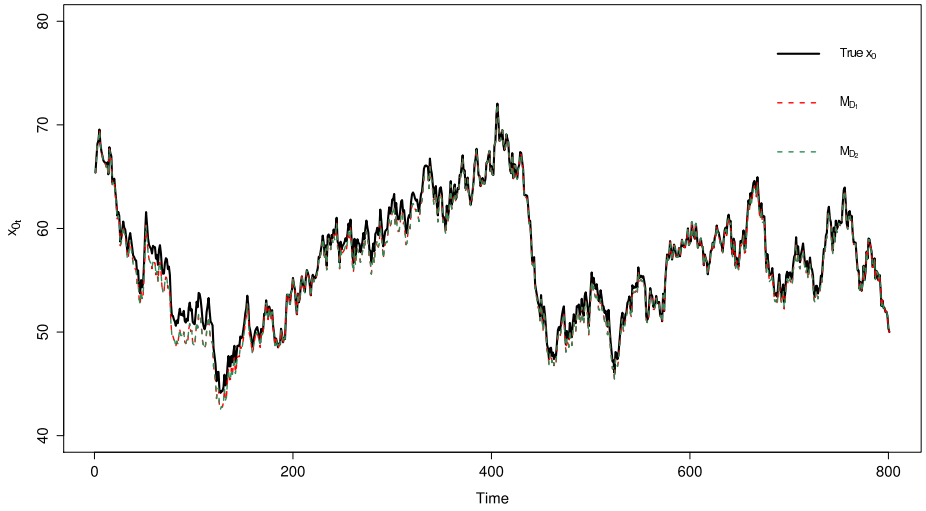


Figure 2.10: *Time series of M_{D1} and M_{D2} estimates in the interpolation case with gain fluctuations $g_t = 0.1sin(0.025t)$ (burn-in = 200).*

Table 2.13: *Comparison of calibration approaches M_{D1} and M_{D2} when interpolating to estimate x_{0t} with sinusoidal gain fluctuation.*

2 References-Sinusoidal Gain- w/Burn In = 200

	Mean Squared Error	Coverage Probability	Interval Width
M_{D1}	0.63553	0.72185	1.15722
M_{D2}	0.63333	0.96380	3.77191

absurd to believe that this would happen in practice because one would test the significance of β_{1t} (Myers 1990; Montgomery *et al.* 2012) for using any method where the possibility of dividing by zero could occur. We demonstrate this by generating data where $\beta_{0t} \approx 2$ for all time and β_{1t} drifts from 1 to -1 over time where $t = 1, \ldots, 1000$ (see Figure 2.11). Figure 2.12 shows the dynamic method is close to the true values of x_{0t} until β_{1t} get close to 0. Within the region where the slope crosses the $x-$axis the posterior estimates become *unstable*. Here we define unstable as meaning that we are within a region where there is division by zero. This instability is only present when $|\beta_{1t}| < \epsilon$, for every $\epsilon \approx 0$. As long as $|\beta_{1t}| > 0$ the dynamic method will perform well when estimating x_{0t}.

Some calibration problems are not linear or approximately linearly related in x_{0t} and y_{0t}. Future work is to investigate the dynamic calibration methods in the presence of nonlinearity. In such settings we may not have the ability to use only 2-points as references. Any approach

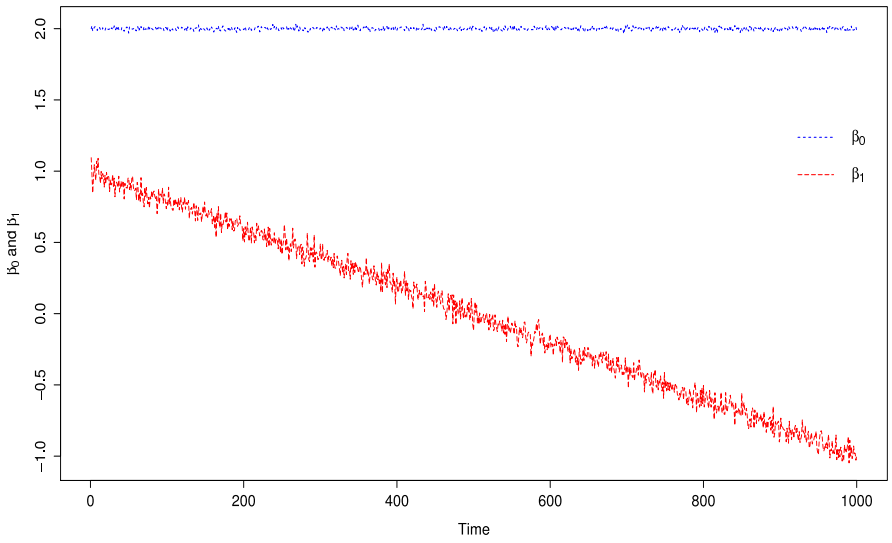


Figure 2.11: *True time series of β_{0t} and β_{1t}*

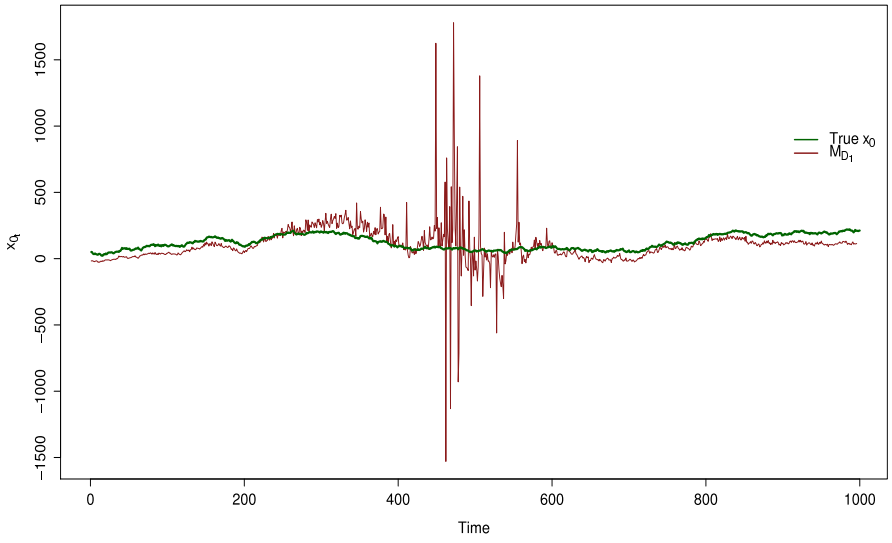


Figure 2.12: *Time series of true x_{0t} and M_{D1} estimate of x_{0t}*

will require more references in order to accurately capture the nonlinear behavior. Another area to be explored is using semiparametric regression, which also allows for parameter variation across time and could be implemented in a near real time setting.

2.6 References

[1] Albert, J. (2007). *Bayesian Computation with R*. New York: Springer.

[2] Bernardo, J. M., and Smith, A. F. M. (1994). *Bayesian Theory*. John Wiley and Sons.

[3] Bremer, J. C. (1979). Improvement of Scanning Radiometer Performance by Digital Reference Averaging. *Instrumentation and Measurement, IEEE Transactions on* 28.1: 46-54.

[4] Brown, P. J. (1993). *Measurement, Regression, and Calibration*. Oxford University Press.

[5] Draper, N. R., and Smith, H. (1998). *Applied Regression Analysis*. 3rd ed., John Wiley, Hoboken, N. J.

[6] Eisenhart, C. (1939). The interpretation of certain regression methods and their use in biological and industrial research. *Annals of Mathematical Statistics*. **10**, 162-186.

[7] Eno, D. R. (1999). Noninformative Prior Bayesian Analysis for Statistical Calibration Problems. Doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

[8] Eno, D. R., and Ye, K. (2000). Bayesian reference prior analysis for polynomial calibration models. *Test*, **9**(1), 191-208.

[9] Givens, G. H., and Hoeting, J. A. (2005). *Computational Statistics*. John Wiley and Sons.

[10] Halperin, M. (1970). On Inverse Estimation in Linear Regression. *Technometrics*. **12**, 727-736.

[11] Hoadley, B. (1970). A Bayesian look at Inverse Linear Regression. *Journal of the American Statistical Association*. **65**, 356-369.

[12] Hunter, W., and Lamboy, W. (1981). A Bayesian Analysis of the Linear Calibration Problem. *Technometrics*. **23**, 323-328.

[13] Imaoka, K., Kachi, M., Kasahara, M., Ito, N., Nakagawa, K., and Oki, T. (2010). Instrument Performance and Calibration of AMSR-E and AMSR2. *International Archives of the Photogrammetry, Remote Sensing and Special Information Science*. **38**(Part 8).

[14] Krutchkoff, R. G. (1967). Classical and Inverse Regression Methods of Calibration. *Technometrics*. **9**, 425-439.

[15] Lwin, T. and Maritz, J. S. (1982). An Analysis of the Linear Calibration Controversy from the Perspective of Compound Estimation. *Technometrics*. **24**, 235-242.

[16] Montgomery, D.C., Peck, E.A. and Vining, G.G. (2012). *Introduction to Linear Regression Analysis*. Wiley.

[17] Myers, R. H. (1990). *Classical and Modern Regression With Applications*. Duxbury/Thompson Learning.

[18] Osborne, C. (1991). Statistical Calibration: A Review. *International Statistical Review*. **59**, 309-336.

[19] Ott, R.L. and Longnecker, M. (2009). *An Introduction to Statistical Methods and Data Analysis*. 6th ed., Brooks/Cole: Cengage Learning.

[20] Özyurt, Ö., and Erar, A. (2003). Multivariate Calibration in Linear Regression and its Application. *Hacettepe Journal of Mathematics and Statistics*. **32**, 53-63.

[21] Petris, G., Petrone, S., and Campagnoli, P. (2009). *Dynamic Linear Models with R*. Springer-Verlag, New York.

[22] R Development Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL: http://www.R-project.org.

[23] Racette, P., Adler, R. F., Wang, J. R., Gasiewski, A. J., and Zacharias, D. S. (1996). An Airborne Millimeter-wave Imaging Radiometer for Cloud, Precipitation, and Atmospheric Water Vapor Studies. *Journal of Atmospheric and Oceanic Technology*, **13**(3), 610-619.

[24] Racette, P. E. (2005). Radiometer Design Analysis Based Upon Measurement Uncertainty (Unpublished doctoral dissertation). The George Washington University, Washington D.C.

[25] Racette, P. E., and Lang, R. H. (2005). Radiometer Design Analysis Based Upon Measurement Uncertainty. *Radio Science*. **40**(5).

[26] Ulaby, F. T., Moore, R. K., and Fung, A. K. (1981). *Microwave Remote Sensing:Active and Passive, Vol. I: Fundamentals and Radiometry*. Addison-Wesley, Advanced Book Program, Reading, Massachusetts.

[27] West, M., Harrison, P. J., and Migon, H. S. (1985). Dynamic generalized linear models and Bayesian forecasting. *Journal of the American Statistical Association*, **80**(389), 73-83.

[28] West, M. and P.J. Harrison (1997). Bayesian Forecasting and Dynamic Linear Models, 2nd Edition. Springer-Verlag, New York.

Chapter 3

Dynamic Bayesian Calibration When the Relationship is Nonlinear

Abstract

Statistical calibration where the relationship between the explanatory variable **X** and response variable **Y** is nonlinear is important in many areas, such as analytical chemistry and radiometry. Especially in radiometry, instrument characteristics change over time, thus calibration is a process that must be conducted as long as the instrument is in use. We propose a dynamic Bayesian method to perform calibration in the presence of a curvilinear relationship between the reference measurements and the response variable. The dynamic calibration approach adequately derives time dependent calibration distributions in the presence of drifting regression parameters. The method is applied to microwave radiometer data and simulated spectroscopy data based on work by Lundberg and de Maré (1980).

3.1 Introduction

In many areas such as analytical chemistry, bioassay, spectroscopy, and radiometry, fitting a curve through data to perform statistical calibration is of great importance. The statistical calibration problem is typically carried out in two stages; first samples are collected consisting of observations and known reference measurements of a targeted subject, and second a fitted curve is established from the first stage and an observed value y_0 is used to predict an unknown targeted reference measurement x_0. The linear approach to this problem has been given much consideration from both the frequentist perspective (Eisenhart 1939; Krutchkoff 1967; Berkson 1969; Williams 1969; Halperin 1970; Martinelle 1970; Lwin and Maritz 1982) and the Bayesian perspective (Hoadley 1970; Hunter and Lamboy 1981; Eno 1999). The multivariate case to the linear calibration problem is considered by Brown (1982) from both

persepectives. Bayesian dynamic approaches to the linear statistical calibration problem have been explored that consider calibration estimates as a function of time (Smith and Corbett 1987; Rivers and Boone 2014).

Unfortunately, in many cases these curves are curvilinear, and straight-line linear methods are inappropriate. Several authors have considered Bayesian nonlinear approaches to the calibration problem. Racine-Poon (1988) used a Bayesian approach to a nonlinear calibration problem arising from agrochemical soil bioassays (Osborne 1991). Racine-Poon (1988) show that the posterior distribution of an unknown concentration η can be calculated by several methods: maximum likelihood; a numerical integration method based on the Gauss quadrature approach of Naylor and Smith (1982); or an approximation based on the Laplace method for integrals (Tierney and Kadane 1986). A noninformative reference prior (Bernardo 1979) approach for the polynomial calibration model is presented by Eno and Ye (2000). Through a second-degree bioassay example presented by Aitchison and Dunsmore (1975), Eno and Ye (2000) derive a reference prior and make posterior inferences about the calibration distribution. In cases when it is not feasible to transform the data to create a straight line, Eno and Ye (2000) show that the inclusion of a quadratic term appropriately adds flexibility to the model. A Bayesian random effects model is proposed by Fong et al. (2012) for the nonlinear calibration problem. Fong et al. (2012) proposed a calibration method that is robust to dependent outliers. They demonstrated the proposed method on data from the HIV Vaccine Trials Network Laboratory and used a normal-mixture model with dependent error terms to model the experimental noise.

Graphite furnace atomic absorption spectroscopy (GFAAS) is an analytical technique for determining trace metal concentrations in different samples. In GFAAS, calibration curves tend to be quadratic and polynomial regression is used to evaluate unknown sample concentrations. Given a known sample of concentrations (x_i) and corresponding absorbance values (y_i) the following nonlinear relationship can be assumed:

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i, \qquad i = 1, 2, \ldots, m, \tag{3.1}$$

where β_0, β_1, β_2 are the unknown parameters and ϵ_i is an independent zero-mean error term with variance σ^2. Also, an observed absorbance y_0 corresponding to an unknown concentration ξ can be modeled as

$$y_0 = \beta_0 + \beta_1 \xi + \beta_2 \xi^2 + \epsilon, \tag{3.2}$$

where the error term again is assumed to be independent with mean of zero and variance σ^2. The estimate of the unknown value ξ is calculated as follow:

$$\xi^* = \frac{-\hat{\beta}_1 + \sqrt{\hat{\beta}_1^2 - 4\hat{\beta}_2\left(\hat{\beta}_0 - y_0\right)}}{2\hat{\beta}_2}, \tag{3.3}$$

where $\hat{\beta}_0$, $\hat{\beta}_1$ and $\hat{\beta}_2$ are the least squares estimates of β_0, β_1 and β_2. The estimate of ξ is denoted as ξ^* and y_0 is the observed value associated with the unknown concentration ξ.

Using Equations (3.1) - (3.3) Lundberg and de Maré (1980) propose a simple interval estimation approach to the spectroscopy calibration problem when there is small measurement error.

For sufficiently small values of the measurement variance σ^2, Lundberg and de Maré (1980) state that there is a unique and consistent estimate ξ^* of ξ and an asymptotic confidence interval of ξ is obtained as

$$\xi^* \pm t_{\alpha/2,(m+n)}\dot{d}(\xi^*)$$

where $d(\xi^*)$ is chosen as

$$d(\xi^*) = \frac{s\left[\frac{1}{n} + \boldsymbol{\Xi}'(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{\Xi}\right]^{1/2}}{\left|\sum_{k=1}^{2} k\hat{\beta}_k\xi^{*k}\right|},$$

where s^2 is the residual variance, $\mathbf{X}$ is the design matrix and

$$\boldsymbol{\Xi} = \begin{bmatrix} 1 \\ \xi^* \\ \xi^{*2} \end{bmatrix}.$$

François et al. (2004) examined optimal designs for linear and nonlinear calibration models. They show that for the quadratic model given by Equation (3.1), the design points should be to the left part of the calibration domain where the calibration curve slope is smaller. François et al. (2004) state that in this area of the domain, the calibration prediction variance is higher and the design points aim then at decreasing the lack of predictive ability of the model in that area. Conversely, the delta method is used to derive an asymptotic confidence interval for ξ by François et al. (2004) as well as Kirkup and Mulholland (2004). The variance in ξ^*, written as $\sigma^2_{\xi^*}$, is given by

$$\sigma^2_{\xi^*} = \left(\frac{\partial\xi^*}{\partial y_0}\sigma_{y_0}\right)^2 + \mathbf{d}'_{\xi^*}\mathbf{V}\mathbf{d}_{\xi^*}, \tag{3.4}$$

where $\mathbf{V}$ is the variance-covariance matrix of $\boldsymbol{\beta}$ and

$$\mathbf{d}_{\xi^*} = \begin{bmatrix} \frac{\partial\xi^*}{\partial\beta_0} \\ \frac{\partial\xi^*}{\partial\beta_1} \\ \frac{\partial\xi^*}{\partial\beta_2} \end{bmatrix},$$

thus deriving an asymptotic confidence interval for ξ

$$\xi^* \pm z_{1-\alpha/2}\sqrt{\sigma^2_{\xi^*}}. \tag{3.5}$$

For an overview of linear and curvilinear calibration methods that commonly use the quadratic calibration model see Merkle (1983), Kirkup and Mulholland (2004), Lavagnini and Magno

(2006), and Lim and Yun (2010).

Weinreb et al. (1990) consider a quadratic relationship in calibration of the advanced very high resolution radiometer (AVHRR). It is assumed that AVHRR can be calibrated by only two points; an internal calibration target (ICT) and cosmic space. Weinreb et al. (1990) state that by not accounting for nonlinearity, errors as large as $2°C$ can be inferred in the estimated scene temperatures. The most direct way to handle the nonlinearity would be to use a quadratic calibration equation (Weinreb et al. 1990).

Calibrations are never concluded once and for all. Instrument characteristics are altered by time and use, especially in radiometry, and calibration must be viewed as an iterative process as long as the instrument is in use (Cervenka and Massa 1994). Our study is motivated by extending the dynamic linear calibration model of Rivers and Boone (2014) to incorporate a quadratic term in the presence of nonlinearity. In Section 3.2, we introduce a Bayesian dynamic nonlinear calibration model akin to that of Lundberg and de Maré (1980), Weinreb et al. (1990), Eno (1999), Eno and Ye (2000), François et al. (2004), Kirkup and Mulholland (2004), Hibbert (2006), Lavagnini and Magno (2006), and Lim and Yun (2010). In Section 3.3, we demonstrate through a simulation study how the dynamic nonlinear calibration model performs alongside the static estimator given by Equation (3.3) under various noise conditions. In Section 3.4 the proposed method is applied to a spectroscopy example and a microwave radiometry example. In the first example, the method is used to determine trace amounts of cadmium (Cd) in water samples and for the second example, it is used to estimate a reference temperature given an observed voltage output measure. In Section 3.5 we conclude with future work and other considerations.

3.2 Dynamic Nonlinear Calibration Model

When collecting laboratory or field measurements for the purpose of calibration, scientists and engineers face a problem when the subsequent stability of the instrument changes in relation to time, temperature, pressure, or some other external factors. These changes may cause the instrument readings to drift since an initial calibration, thus making it necessary to recalibrate the instrument (Ziemer and Strauss, 1978). We address this problem by developing a dynamic calibration approach that detects changes in the calibration constants in the presence of a curvilinear relationship.

Let $\{(\mathbf{X}, \mathbf{Y}_t)|t = 1, 2, \ldots, T\}$ be the reference measurement and responses in the calibration experiment at time t and suppose the relationship can be described by

$$\mathbf{Y}_t = \mathbf{X}\boldsymbol{\beta}_t + \boldsymbol{\epsilon}_t, \qquad t = 1, 2, \ldots, T, \tag{3.6}$$

where $\mathbf{Y}_t$ is a series of $r-$dimensional ($r \equiv n$) vector of responses, $\mathbf{X}$ is the fixed ($r \times d$) ($d = p + 1$) reference design matrix,

$$\begin{bmatrix} 1 & X_1 & X_1^2 \\ 1 & X_2 & X_2^2 \\ \vdots & \vdots & \vdots \\ 1 & X_r & X_r^2 \end{bmatrix},$$

$\boldsymbol{\beta}_t = [\beta_{0t}\ \beta_{1t}\ \ldots\ \beta_{dt}]'$, is a series of $d-$dimensional vectors of unknown dynamic regression parameters, and $\boldsymbol{\epsilon}_t$ is a $r-$dimensional vector of independently normally distributed error terms with mean $\mathbf{0}$ and variance-covariance matrix $\mathbf{E} = \sigma_E^2\mathbf{I}$. Equation (3.6) is known as the observation equation.

The evolving relationship between $\mathbf{X}$ and $\mathbf{Y}_t$ is expressed by the dynamic parameter vector $\boldsymbol{\beta}_t$. The evolution in time of the regression parameters is modeled as

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\omega}_t, \qquad t = 1, 2, \ldots, T, \tag{3.7}$$

where $\boldsymbol{\omega}_t$ is a $d-$dimensional vector of independently normally distributed error terms with mean $\mathbf{0}$ and variance-covariance matrix $\mathbf{W} = \sigma_W^2\boldsymbol{\Omega}$, and $\boldsymbol{\Omega} = [\mathbf{X}'\mathbf{X}]^{-1}$. Equation (3.7) is known as the system equation.

The observation vector $\mathbf{Y}_t$ and the dynamic regression parameter vector $\boldsymbol{\beta}_t$ are both random variables, thus the expected values $\hat{\mathbf{Y}}_t$ and $\hat{\boldsymbol{\beta}}_t$ are the means of their respective distributions and must be estimated sequentially. The one-step forecast for $\mathbf{Y}_t$ and posterior distributions $\boldsymbol{\beta}_t$ for each time t are as follows in Algorithm 2 (Dynamic Linear Regression Models algorithm). See West et al. (1985), or West and Harrison (1997) for a more detailed discussion of Dynamic Linear Regression Models (DLRMs).

Furthermore, let $\{y_{0t}|t = 1, 2, \ldots, T\}$ be the observation from the second stage of the calibration experiment corresponding to an unknown reference of interest x_{0t}, and

$$y_{0t} = \beta_{0t} + \beta_{1t}x_{0t} + \beta_{2t}x_{0t}^2 + \epsilon_{0t}, \qquad t = 1, 2, \ldots, T, \tag{3.8}$$

where β_{0t}, β_{1t}, and β_{2t} are the time dependent regression coefficients and ϵ_{0t} are assumed to be independently normally distributed with mean 0 and variance σ_E^2. Since the quadratic model in Equations (3.6) and (3.7) are not monotonic on $\mathbb{R}$, the domain will be restricted to where it is strictly increasing.

We assume that the first stage of the calibration experiment is independent of the second stage, therefore x_{0t} is independent of $(\boldsymbol{\beta}_t, \boldsymbol{\Gamma})$ where the joint prior distribution is

$$\pi(x_{0t}, \boldsymbol{\beta}_t, \boldsymbol{\Gamma}) = \pi(x_{0t})\pi(\boldsymbol{\beta}_t, \boldsymbol{\Gamma}) \tag{3.9}$$

and $\boldsymbol{\Gamma}' = [\sigma_E^2\ \ \sigma_W^2]$. The posterior of $(x_{0t}, \boldsymbol{\beta}_t, \boldsymbol{\Gamma})$ is then given by

$$\begin{aligned} \pi(x_{0t}, \boldsymbol{\beta}_t, \boldsymbol{\Gamma}|y_{0t}, \mathbf{Y}_t) &\propto f(y_{0t}, \mathbf{Y}_t|x_{0t}, \boldsymbol{\beta}_t, \boldsymbol{\Gamma})\pi(x_{0t}, \boldsymbol{\beta}_t, \boldsymbol{\Gamma}) \\ &\propto f(y_{0t}|x_{0t}, \boldsymbol{\beta}_t, \boldsymbol{\Gamma})\pi(x_{0t})f(\mathbf{Y}_t|\boldsymbol{\beta}_t, \boldsymbol{\Gamma})\pi(\boldsymbol{\beta}_t, \boldsymbol{\Gamma}) \\ &\propto f(y_{0t}|x_{0t}, \boldsymbol{\beta}_t, \boldsymbol{\Gamma})\pi(\boldsymbol{\beta}_t, \boldsymbol{\Gamma}|\mathbf{Y}_t)\pi(x_{0t}), \end{aligned} \tag{3.10}$$

Algorithm 2: Updating (*DLRM*) Dynamic Linear Regression Model

Initialize $t = 0$
{Initial information $(\boldsymbol{\beta}_0|D_0) \sim N_d[\mathbf{m_0}, \mathbf{C_0}]$}
Input: $\mathbf{m_0}$, $\mathbf{C_0}$, $\mathbf{E}$, $\mathbf{W}$
loop
  $t = t + 1$
  {Compute prior at t: $(\boldsymbol{\beta}_t|D_{t-1}) \sim N_d[\mathbf{a_t}, \mathbf{R_t}]$}
    $\mathbf{a}_t = \mathbf{m}_{t-1}$
    $\mathbf{R}_t = \mathbf{C}_{t-1} + \mathbf{W}$
  Input: $\mathbf{X}$
  {Compute forecast at t: $(\mathbf{Y_t}|D_{t-1}) \sim N_r[\mathbf{f_t}, \mathbf{Q_t}]$}
    $\mathbf{f}_t = \mathbf{X}\mathbf{a}_t$
    $\mathbf{Q}_t = \mathbf{X}\mathbf{R}_t\mathbf{X}' + \mathbf{E}$
  Input: $\mathbf{Y}_t$
  {Compute forecast error $\mathbf{e}_t$}
    $\mathbf{e}_t = \mathbf{Y}_t - \mathbf{f}_t$
  {Compute adaptive gain matrix $\mathbf{A}_t$}
    $\mathbf{A}_t = \mathbf{Q}_t^{-1}\mathbf{X}\mathbf{R}_t$
  {Compute posterior at t: $(\boldsymbol{\beta}_t|D_t) \sim N_d[\mathbf{m}_t, \mathbf{C}_t]$}
    $\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_t\mathbf{e}_t$
    $\mathbf{C}_t = \mathbf{R}_t - \mathbf{A}_t'\mathbf{Q}_t\mathbf{A}_t$
end loop

where $\mathbf{Y}_t$ and y_{0t} are respectively the observations from the first and second stages of calibration. Our knowledge about the evolving relationship established in the calibration experiment at each time point is given by the posterior density $\pi(\boldsymbol{\beta}_t, \boldsymbol{\Gamma}|\mathbf{Y}_t)$ which is the middle term in Equation (3.10) and found by Algorithm 2 for given values of $\boldsymbol{\Gamma}$. Following Algorithm 2 and using multivariate normal theory (West and Harrison 1997) we have

$$\pi(\boldsymbol{\beta}_t|\mathbf{Y}_t, \boldsymbol{\Gamma}) \sim N_d[\mathbf{m}_t, \mathbf{C}_t],$$

where $\mathbf{m}_t$ is the posterior mean and $\mathbf{C}_t$ is the variance-covariance matrix of $\boldsymbol{\beta}_t$ at time t.

The first term in Equation (3.10), $f(y_{0t}|x_{0t}, \boldsymbol{\beta}_t, \boldsymbol{\Gamma})$, is the likelihood function for the second stage of the calibration experiment which provides information from the data, and $\pi(x_{0t})$ is the prior density for the unknown calibration reference x_{0t}. We wish to obtain the conditional posterior density $\pi(x_{0t}|y_{0t}, \mathbf{Y}_t)$ at each time t. In order to achieve this we will have to integrate over $(\boldsymbol{\beta}, \boldsymbol{\Gamma})$.

We reduce the parameter space by centering and scaling the data such that

$$\sum_{i=1}^{r} x_i = 0 \text{ and } \frac{1}{n}\sum_{i=1}^{r} x_i^2 = 1$$

and the intercept term β_{0t} is eliminated by moving the origin of the calibration (Hibbert 2006) so the model in Equation (3.8) is written as

$$y_{0t} - \bar{y}_t = \beta_{1t}\left(x_{0t} - \bar{x}\right) - \beta_{2t}\left(x_{0t}^2 - \overline{x^2}\right) + \epsilon_{0t}, \qquad t = 1, 2, \ldots, T, \tag{3.11}$$

where $\bar{x} = \frac{1}{r}\sum_{i=1}^{r} x_i$ and $\overline{x^2} = \frac{1}{r}\sum_{i=1}^{r} x_i^2$. In the second stage of the calibration experiment, an observation y_{0t} allows calculation of an unknown reference at time t by:

$$\hat{x}_{0t} = \frac{-\hat{\beta}_{1t} \pm \sqrt{\hat{\beta}_{1t}^2 - 4\hat{\beta}_{2t}\left(\bar{y}_t - y_{0t} - \hat{\beta}_{1t}\bar{x} - \hat{\beta}_{2t}\overline{x^2}\right)}}{2\hat{\beta}_{2t}}. \tag{3.12}$$

The quadratic model in Equation (3.12) has two possible roots on $\mathbb{R}$ but the solution of interest depends on the sign of $\hat{\beta}_{1t}$. The solution is the increasing part of Equation (3.12) when $\hat{\beta}_{1t} < 0$ and the decreasing part of Equation (3.12) is the solution when $\hat{\beta}_{1t} > 0$.

Given the reduced parameter space and Equation (3.11), the likelihood function at time t is expressed as

$$f(y_{0t}|x_{0t}, \boldsymbol{\beta}_t, \boldsymbol{\Gamma}) \propto \exp\left\{ -\frac{1}{2}\left[\sigma_{Y_t}^{-2}(y_{0t} - \bar{y}_t - \beta_{1t}(x_{0t} - \bar{x}) - \beta_{2t}(x_{0t}^2 - \overline{x^2}))^2\right] \right\},$$

the prior density at time t is

$$\pi(x_{0t}) \propto \exp\left\{ -\frac{x_{0t}^2}{2} \right\},$$

and integrating Equation (3.10) with respect to $\boldsymbol{\beta}_t$ for a given $\boldsymbol{\Gamma}$, produces

$$\begin{aligned}
\pi(x_{0t}|y_{0t}, \mathbf{Y}_t, \boldsymbol{\Gamma}) &\propto f(y_{0t}|x_{0t}, \boldsymbol{\Gamma}, \mathbf{Y}_t)\pi(x_{0t}) \\
&\propto \exp\left\{ -\frac{1}{2}\left[\sigma_{Y_t}^{-2}(y_{0t} - \bar{y}_t - \beta_{1t}(x_{0t} - \bar{x}) - \beta_{2t}(x_{0t}^2 - \overline{x^2}))^2\right] - \frac{1}{2}x_{0t}^2 \right\} \\
&\propto \exp\left\{ -\frac{1}{2}\left[\sigma_{Y_t}^{-2}(\hat{x}_{0t} - x_{0t})^2\right] - \frac{1}{2}x_{0t}^2 \right\}.
\end{aligned}$$

By completing the square, the posterior density for the unknown reference measurement at time t is

$$\pi(x_{0t}|y_{0t}, \mathbf{Y}_t, \boldsymbol{\Gamma}) \sim N[\mu_{x_{0t}}, \sigma_{x_{0t}}^2], \tag{3.13}$$

with

$$\begin{aligned}
\mu_{x_{0t}} &= \frac{x_{0t}}{1 + \sigma_{Y_t}^2}, \\
\sigma_{x_{0t}}^2 &= \frac{1}{1 + \sigma_{Y_t}^2},
\end{aligned}$$

and

$$\sigma^2_{Y_t} = \text{tr}(\mathbf{Q}_t)$$

where tr($\cdot$) denotes trace of the one-step forecast variance-covariance matrix $\mathbf{Q}_t$ which is defined in Algorithm 2.

To obtain posterior samples of Equation (3.13) for a given $\mathbf{\Gamma}$ we use the Sampling Importance Resampling (SIR) algorithm (Rubin 1987; Smith and Gelfand 1992; Givens and Hoeting 2005; Albert 2007) to draw random samples from the target distribution $\pi(x_{0t}|y_{0t}, \mathbf{Y}_t, \mathbf{\Gamma})$ by using a candidate distribution $g(\mathbf{\Theta})$ for $\mathbf{\Gamma}$, thus deriving the posterior densities of interest. Algorithm 3 is the Sampling Importance Resampling algorithm.

Algorithm 3: (*SIR*) Sampling Importance Resampling

---

1. Draw **sample** candidates $(\mathbf{\Theta}^{(1)}), \ldots, (\mathbf{\Theta}^{(m)})$ *i.i.d.* from $g(\mathbf{\Theta})$

2. Calculate the standardized **importance** weights,
   $w(\mathbf{\Theta}^{(i)}) = \frac{f(\mathbf{\Theta}^{(i)})/g(\mathbf{\Theta}^{(i)})}{\sum f(\mathbf{\Theta}^{(i)})/g(\mathbf{\Theta}^{(i)})}$ for $i = 1, \ldots, m$

3. **Resample** $\mathbf{\Gamma}^{(1)}, \ldots, \mathbf{\Gamma}^{(n)}$ from $\mathbf{\Theta}^{(1)}, \ldots, \mathbf{\Theta}^{(m)}$ with replacement of probability $w(\mathbf{\Theta}^{(1)}), \ldots, w(\mathbf{\Theta}^{(m)})$ respectively.

For most applications, it is believed *a priori* that the observational variance σ^2_E is greater in magnitude than the system variance σ^2_W, such that

$$\sigma^2_W < \sigma^2_E.$$

To enforce this belief about the variance relationship we utilize the following prior distributions:

$$\sigma^2_E \sim Uniform(0, \alpha_E) \tag{3.14}$$

$$\sigma^2_W|\sigma^2_E \sim Uniform(0, \sigma^2_E), \tag{3.15}$$

where prior distributions (3.14) and (3.15) ensure the system variance to be less than the observation variance. Since these are proper prior distributions the resulting posterior distribution will also be proper. We combine prior distributions (3.14), (3.15), Algorithm 2 (DLM), Equation (3.13), and Algorithm 3 (SIR) together and propose the Dynamic Calibration Method in Algorithm 4.

Algorithm 4: Dynamic Calibration

1. Draw M *i.i.d.* **sample** candidates for (σ_E^2, σ_W^2) from $g(\mathbf{\Theta})$.

2. Calibration data are fit using the DLRM framework for each of the M proposal samples $(\sigma_E^{2(m)}, \sigma_W^{2(m)})$, with the prior moments for $(\boldsymbol{\beta}_0|D_0)$ as $\mathbf{m_0} = \mathbf{1_d}$ and $\mathbf{C_0} = \mathbf{100I_{(d\times d)}}$, where $\mathbf{1_d}$ is a $d-$dimensional vector of ones:

   a. Data are scaled and centered such that $\sum_{i=1}^{r} x_i = 0$, $\frac{1}{n}\sum_{i=1}^{r} x_i^2 = 1$ and $\beta_{0t} = 0$;

   b. Estimate $\boldsymbol{\beta}_t^{(m)}|\sigma_E^{2(m)}, \sigma_W^{2(m)}$ for the m^{th} proposal sample $\mathbf{\Theta}^{(m)}$ for all time t;

   c. Draw a sample from $\pi(x_{0t}|y_{0t}, \mathbf{Y}_t, \mathbf{\Gamma})$ given the m^{th} proposal sample $\mathbf{\Theta}^{(m)}$ at time t;

   d. Calculate log-likelihoods, $log[f(\mathbf{\Theta}^{(m)})]$, for each $(\sigma_{E_t}^{2(m)}, \sigma_{W_t}^{2(m)})$ pair.

3. Sampling Importance Resampling (SIR) is used to simulate samples of $x_{0t}|\boldsymbol{\beta}_t, \sigma_E^2, \sigma_W^2$ by accepting a subset of N from the proposal density to be distributed according to the posterior density $\pi(\mathbf{\Gamma}|\mathbf{Y}_t)$ with candidate density $g(\mathbf{\Theta})$.

   a. Calculate the standardized **importance** weights, $w(\mathbf{\Theta}^{(1)}), \ldots, w(\mathbf{\Theta}^{(M)})$, where $w(\mathbf{\Theta}^{(m)}) = log[f(\mathbf{\Theta}^{(m)})] - log[g(\mathbf{\Theta}^{(m)})]$ for the m^{th} proposal sample;

   b. **Resample** N calibrated time series from the M proposal values with replacement given probabilities $p(\mathbf{\Theta}^{(m)})$ where

$$p(\mathbf{\Theta}^{(m)}) = \frac{e^{w(\mathbf{\Theta}^{(m)})}}{\sum_{j=1}^{M} e^{w(\mathbf{\Theta}^{(j)})}}.$$

4. Rescale calibrated time series to original scale and record summary statistics (i.e. medians and credible sets) across each time t .

3.3 Simulation Study

In this section, several simulation experiments are conducted to illustrate the dynamic behavior of the proposed calibration method. The simulation study considers three possible system scenarios. The first simulation design is thought to be a perfect system that is without random fluctuations or shocks. In other words, the variability in observed time series is due to random noise. Our second simulation is designed to give insight into the behavior of the dynamic approach when there is a random stepped shock experienced in the system (see Figure 3.1). This random shock is imposed on the slope parameter β_{1t}. The third simulation design is more drastic than the second design. We force a sinusoidal disturbance on the slope parameter β_{1t}. This design creates a scenario where the model is disturbed for some longer period of time than just a short random shock (see Figure 3.1). For the second and third simulation scenarios we model the dynamic system in vertex form to help illustrate

how fluctuations affect the quadratic model,

$$y_t = \hat{s}_t(x - \hat{h}_t)^2 + \hat{k}_t \tag{3.16}$$

where

$$\begin{aligned} \hat{s}_t &= \hat{\beta}_{2t}\gamma_t, \\ \hat{h}_t &= \frac{-\hat{\beta}_{1t}}{2\hat{\beta}_{2t}}, \\ \hat{k}_t &= \hat{\beta}_{0t} - \frac{-\hat{\beta}_{1t}^2}{4\hat{\beta}_{2t}}, \end{aligned}$$

and the multiplicative disturbances in Figure 3.1 are denoted as γ_t. Under each of the three scenarios we consider nine cases based on variance combinations for the observation and system variances, (σ_E^2, σ_W^2), for the Dynamic Linear Models.

The true values for σ_E^2 and σ_W^2 used in the simulation study are $(0.00001, 0.0001, 0.001)$

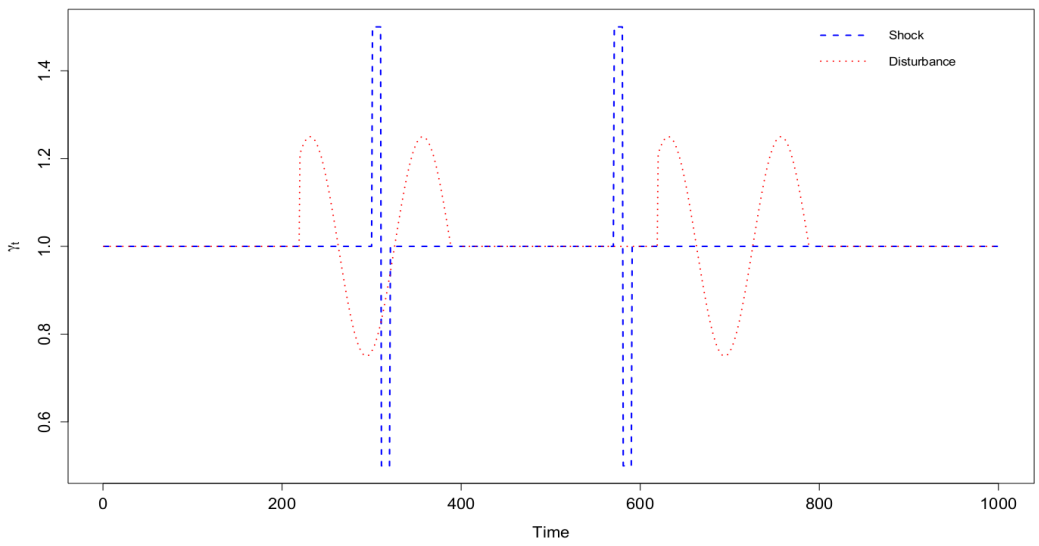


Figure 3.1: *γ_t: Short stepped shock (blue) and long sinusoidal disturbance (red)*

and $(0.00005, 0.0001, 0.001)$, respectively. For each variance pair (σ_E^2, σ_W^2), the number of simulated realizations is $N = 100$. In each realization, the number of simulated time periods is $T = 1000$. The posterior densities samples drawn from $\pi(x_{0t}|y_{0t}, \mathbf{Y}_t, \mathbf{\Gamma})$ by the Dynamic Calibration Approach (Algorithm 4) are assessed via the square root of the averaged mean squared errors ($RAMSE$), the average interval width ($AvIW$), and the average coverage

probability ($AvCP$):

$$RAMSE = \left[\frac{1}{N}\sum_{j=1}^{N} MSE_j\right]^{\frac{1}{2}},$$

where

$$MSE_j = \frac{1}{T}\sum_{t=1}^{T}(\hat{x}_{0t} - x_{0t})^2;$$

and

$$AvIW = \frac{1}{N}\sum_{j=1}^{N} IW_j.$$

where

$$IW_j = \frac{1}{T}\sum_{t=1}^{T}(x_{0t}^U - x_{0t}^L).$$

Note that if x_{0t}^L is the 0.025 posterior quantile for x_{0t}, and x_{0t}^U is the 0.975 posterior quantile for x_{0t}, where x_{0t} is the true value of the calibration target from the second stage of experimentation, then (x_{0t}^L, x_{0t}^U) is a 95% credible interval.

Using the credible interval above we defined the coverage probability (CP) which is calculated as such

$$CP_j = \frac{1}{T}\sum_{t=1}^{T}\psi_t$$

where

$$\psi_t = P[x_{0t}^L < x_{0t} < x_{0t}^U] = \begin{cases} 0 & \text{if } x_{0t} \notin (x_{0t}^L, x_{0t}^U); \\ 1 & \text{if } x_{0t} \in (x_{0t}^L, x_{0t}^U). \end{cases}$$

The average coverage probability ($AvCP$) is calculated by averaging across the number of replications in the simulation study, where

$$AvCP = \frac{1}{N}\sum_{j=1}^{N} CP_j.$$

The data are generated from the following model:

$$\mathbf{Y}_t = \mathbf{X}\boldsymbol{\beta}_t + \boldsymbol{\epsilon}_t, \tag{3.17}$$

where $\mathbf{X}$ is a known fixed model matrix of reference values augmented with a column of 1's. The dynamic Bayesian nonlinear calibration model will be assessed across three different reference measurement schemes. In the first scheme, the reference measurements will be taken at [20, 90, 100]. The references will be placed at [20, 60, 90, 100] for the second scheme and at [20, 40, 60, 90, 100] for the final and third scheme. See Figure 3.2.

The vector of regression parameters, $\boldsymbol{\beta}_t$, are random draws from a multivariate normal distribution with mean vector $[\beta_0\ \beta_1\ \beta_2]'$ and variance-covariance matrix, $\mathbf{W} = \sigma_W^2 \left[\mathbf{X}'\mathbf{X}\right]^{-1}$ for $t = 1, \ldots, T$, where $\beta_0 = -0.0007$, $\beta_1 = 0.01858$, and $\beta_1 = -0.000117$. For each t, the random multivariate error vector is

$$\boldsymbol{\epsilon}_t \sim N_r[\mathbf{0}, \sigma_E^2 \mathbf{I}]. \tag{3.18}$$

The results of the simulation study are summarized in Tables 3.1 - 3.3. The results summarize the dynamic nonlinear calibration algorithm under different variance pairs, (σ_E^2, σ_W^2), and results for the static model given by Equation (3.3). An asymptotic variance and 95% confidence interval for the static quadratic calibration is calculated from Equations (3.4) and (3.5). The dynamic calibration approach is denoted as DC and the static calibration approach is denoted as SC in the tables.

From the tables, we can observe that when only 3 reference measurements are taken at time t the static method is unable to derive an estimate for σ^2, thus confidence intervals are not calculated. The lack of confidence intervals for the estimate when using the static quadratic model is represented by the (—) for the average interval width (AIW) and the average coverage probability (ACP). In Tables 3.1 - 3.3 we see that the RAMSE for the dynamic method is consistently smaller than the the RAMSE for the static method. As the observation variance σ_E^2 increases the RAMSEs of the dynamic estimator and the static estimator increase for the first and second simulations. It is evident that when the system is undisturbed, additional reference measurements do not improve estimation. The second and third simulations reveals that additional references improve estimation when the system experiences random shocks or some sinusoidal type disturbance. The average interval width (AIW) of the dynamic calibration approach increases as the observational variance increases, thus the average coverage probability (ACP) goes up. The interval width for the static model is roughly always 3.92 and is incapable of adapting to noisy environments. In the sinusoidal scenario summarized in Table 3.3, the RAMSE for the dynamic method decreases as σ_E^2 increases but the inverse occurred with the static method. As expected, the dynamic method performs well in noisier environments than the static method.

Table 3.1: *Main results of the simulation study for the case without a system fluctuation: RAMSE, AIW and AvCP are given the dynamic nonlinear calibration estimator (DC) and the static quadratic estimator (SC) under* (σ_E^2, σ_W^2)*. Results are provided for the 3 reference, 4 reference and the 5 reference models.*

		3 Reference Model						4 Reference Model						5 Reference Model					
		RAMSE		**AIW**		**ACP**		**RAMSE**		**AIW**		**ACP**		**RAMSE**		**AIW**		**ACP**	
σ_W^2	σ_E^2	DC	SC	DC	SC	DC	SC	DC	SC	DC	SC	DC	SC	DC	SC	DC	SC	DC	SC
0.00005	0.00001	0.249	0.382	1.471	—	0.992	—	0.247	0.358	1.500	3.763	0.993	0.960	0.249	0.350	1.569	3.920	0.994	1.000
	0.0001	0.643	1.099	4.642	—	0.997	—	0.739	1.072	4.741	3.920	0.994	0.931	0.768	1.083	4.955	3.920	0.995	0.927
	0.001	1.988	3.391	14.592	—	0.997	—	2.332	3.376	14.875	3.920	0.995	0.451	2.429	3.374	15.515	3.920	0.995	0.452
0.0001	0.00001	0.291	0.413	1.473	—	0.984	—	0.259	0.365	1.500	3.763	0.991	0.960	0.257	0.356	1.569	3.920	0.993	1.000
	0.0001	0.661	1.109	4.645	—	0.996	—	0.747	1.081	4.742	3.920	0.994	0.930	0.770	1.082	4.955	3.920	0.995	0.929
	0.001	1.999	3.399	14.592	—	0.997	—	2.336	3.379	14.875	3.920	0.994	0.451	2.432	3.377	15.515	3.878	0.995	0.447
0.001	0.00001	0.716	0.806	1.500	—	0.688	—	0.427	0.498	1.502	3.920	0.917	0.998	0.370	0.445	1.570	3.920	0.958	0.999
	0.0001	0.936	1.306	4.652	—	0.983	—	0.825	1.131	4.742	3.920	0.990	0.924	0.821	1.120	4.956	3.920	0.993	0.924
	0.001	2.126	3.475	14.600	—	0.996	—	2.380	3.411	14.876	3.920	0.994	0.446	2.462	3.404	15.516	3.920	0.994	0.448

Note: Static method is unable to estimate σ^2 with 3 reference points. Therefore confidence intervals were not calculated.

Table 3.2: *Main results of the simulation study for the case with a systematic shock: RAMSE, AIW and AvCP are given the dynamic nonlinear calibration estimator (DC) and the static quadratic estimator (SC) under* (σ_E^2, σ_W^2)*. Results are provided for the 3 reference, 4 reference and the 5 reference models.*

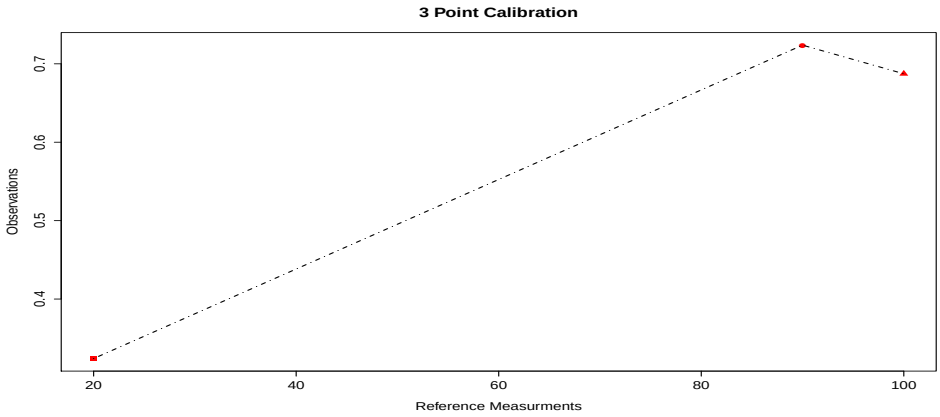
		3 Reference Model						4 Reference Model						5 Reference Model					
		RAMSE		**AIW**		**ACP**		**RAMSE**		**AIW**		**ACP**		**RAMSE**		**AIW**		**ACP**	
σ_W^2	σ_E^2	DC	SC	DC	SC	DC	SC	DC	SC	DC	SC	DC	SC	DC	SC	DC	SC	DC	SC
0.00005	0.00001	2.554	2.603	1.643	—	0.953	—	2.506	2.604	1.641	3.920	0.953	0.960	2.490	2.605	1.715	3.920	0.954	0.960
	0.0001	2.587	2.788	4.770	—	0.958	—	2.562	2.789	4.844	3.920	0.956	0.889	2.553	2.789	5.060	3.920	0.956	0.889
	0.001	3.115	4.234	14.652	—	0.966	—	3.301	4.231	14.921	3.920	0.967	0.429	3.353	4.231	15.560	3.920	0.970	0.429
0.0001	0.00001	2.556	2.605	1.644	—	0.945	—	2.507	2.605	1.640	3.920	0.951	0.960	2.491	2.605	1.715	3.920	0.953	0.960
	0.0001	2.589	2.790	4.770	—	0.958	—	2.563	2.790	4.844	3.920	0.956	0.889	2.554	2.789	5.060	3.920	0.956	0.889
	0.001	3.120	4.239	14.653	—	0.966	—	3.303	4.234	14.921	3.920	0.967	0.429	3.354	4.232	15.560	3.920	0.969	0.429
0.001	0.00001	2.619	2.674	1.672	—	0.674	—	2.524	2.621	1.642	3.920	0.880	0.959	2.502	2.615	1.715	3.920	0.919	0.959
	0.0001	2.654	2.855	4.778	—	0.945	—	2.582	2.807	4.845	3.920	0.952	0.882	2.566	2.800	5.061	3.920	0.954	0.885
	0.001	3.192	4.296	14.656	—	0.962	—	3.330	4.259	14.922	3.920	0.966	0.425	3.373	4.252	15.561	3.920	0.968	0.426

Note: Static method is unable to estimate σ^2 with 3 reference points. Therefore confidence intervals were not calculated.

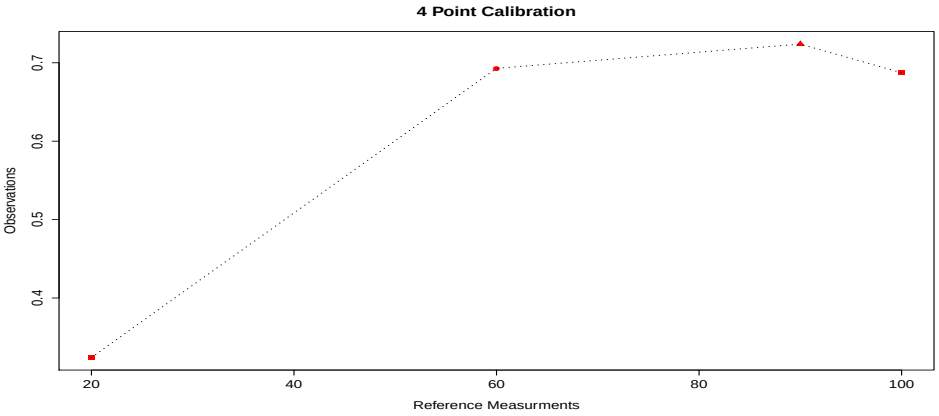
Table 3.3: *Main results of the simulation study for the case with a sinusoidal disturbance: $RAMSE$, AIW and $AvCP$ are given the dynamic nonlinear calibration estimator (DC) and the static quadratic estimator (SC) under (σ_E^2, σ_W^2). Results are provided for the 3 reference, 4 reference and the 5 reference models.*

		3 Reference Model						**4 Reference Model**						**5 Reference Model**					
		RAMSE		**AIW**		**ACP**		**RAMSE**		**AIW**		**ACP**		**RAMSE**		**AIW**		**ACP**	
σ_W^2	σ_E^2	DC	SC	DC	SC	DC	SC	DC	SC	DC	SC	DC	SC	DC	SC	DC	SC	DC	SC
0.00005	0.00001	4.835	4.755	1.596	—	0.648	—	4.782	4.745	1.582	3.920	0.650	0.687	2.189	2.185	1.595	3.920	0.670	0.741
	0.0001	4.886	4.888	4.708	—	0.689	—	2.303	2.425	4.752	3.920	0.762	0.707	2.310	2.422	4.966	3.920	0.769	0.707
	0.001	5.308	5.930	14.626	—	0.856	—	3.210	4.036	14.879	3.920	0.967	0.396	3.280	4.033	15.519	3.920	0.969	0.396
0.0001	0.00001	4.841	4.762	1.597	—	0.643	—	4.783	4.747	1.582	3.920	0.649	0.687	2.192	2.188	1.595	3.920	0.669	0.740
	0.0001	4.892	4.895	4.708	—	0.688	—	2.307	2.429	4.752	3.920	0.760	0.706	2.313	2.426	4.966	3.920	0.768	0.707
	0.001	5.315	5.938	14.626	—	0.855	—	3.215	4.041	14.879	3.920	0.967	0.395	3.283	4.037	15.519	3.920	0.969	0.396
0.001	0.00001	4.914	4.843	1.619	—	0.454	—	2.245	2.240	1.528	3.920	0.630	0.732	2.225	2.220	1.597	3.920	0.651	0.735
	0.0001	4.964	4.974	4.714	—	0.679	—	2.355	2.475	4.753	3.920	0.750	0.698	2.346	2.456	4.967	3.920	0.760	0.700
	0.001	5.390	6.027	14.629	—	0.844	—	3.262	4.081	14.880	3.920	0.964	0.392	3.316	4.069	15.520	3.920	0.968	0.392

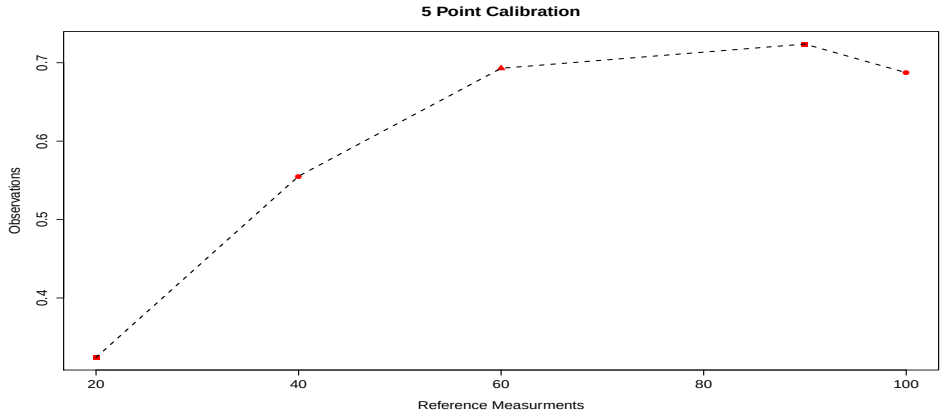
Note: Static method is unable to estimate σ^2 with 3 reference points. Therefore confidence intervals were not calculated.



(a) *3-Point Calibration*



(b) *4-Point Calibration*



(c) *5-Point Calibration*

Figure 3.2: *Three reference design schemes: (a) 3-point calibration; (b) 4-point calibration; (c) 5-point calibration*

3.4 Applications

3.4.1 Example 1

As a demonstration of the dynamic nonlinear calibration method, we extend the example of nonlinear calibration presented by Lundberg and de Maré (1980). In Lundberg and de

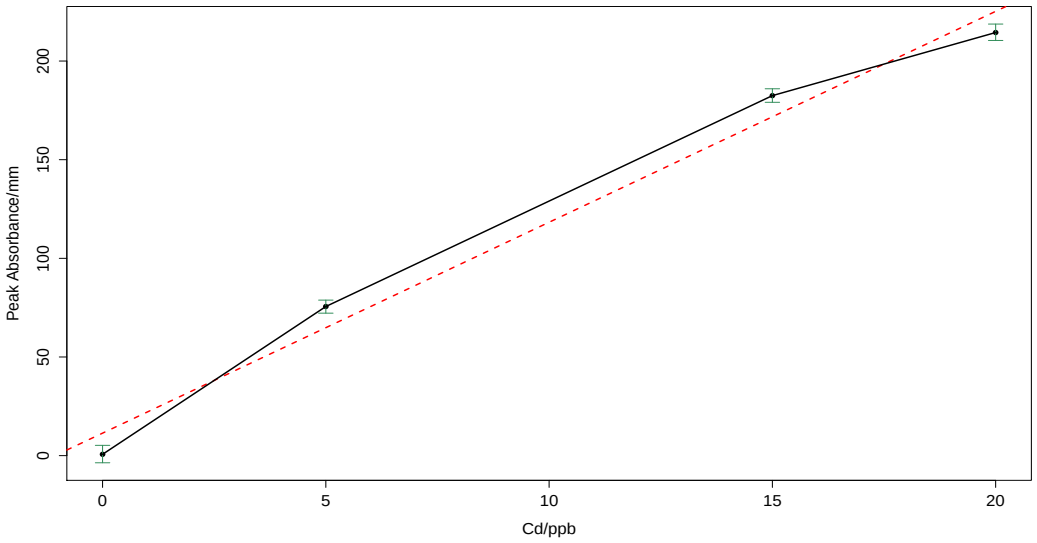


Figure 3.3: *Linear (dotted) and quadratic calibration curves for Cd using four standard references. Indicated are 95% credible intervals of the estimated absorbances.*

Maré's (1980) example, graphite furnace atomic absorption spectroscopy (GFAAS) was used to determine trace amounts of cadmium (Cd) in water samples. In the first stage of experimentation two μl volumes of standard solutions containing $0, 5, 10, 15$, and 20 ppb (parts per billion) of Cd were injected into a graphite furnace (Varian Techtron AA-6 spectrophotometer supplied with a Carbon Rod Atomizer model 63). The transient absorbance signals

Table 3.4: *Cd signals obtained when atomizing standards.*

Concentration/ppb	Peak Absorbance/mm					
0	0	1	1	0	1	
5	74	74	78	78	76	
15	183	184	178	183	184	
20	217	215	213	218	210	215

obtained when atomizing the standards were recorded with a strip chart recorder. Each

standard was run several times and a plot of peak absorbance (in mm) vs. concentration was made using the data in Table 3.4 (See Figure 3.3). At the second stage of experimentation the fifth reference (10 ppb) was used as an unknown sample to test the validity of the method with peak absorbance measurements of 135,142,132, 141, and 136. Lundberg and de Maré (1980) report an approximate 95% confidence interval for the unknown concentration with 10 being the true value as $[9.7, 10.3]$.

We extend Lundberg and de Maré's (1980) example by generating 500 simulated peak

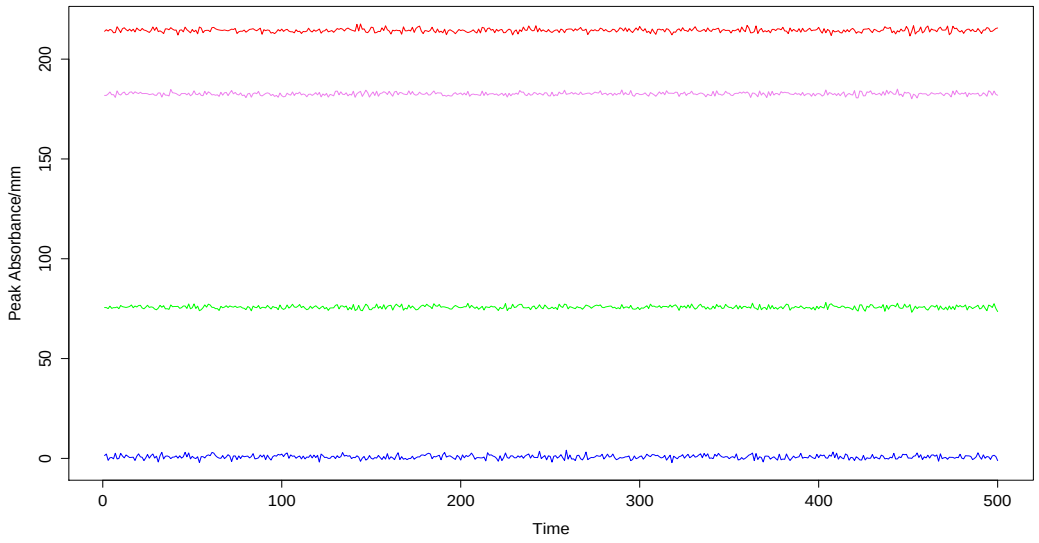


Figure 3.4: *Peak Absorbance measurements: Concentration 0ppb (blue); Concentration 5ppb (green); Concentration 15ppb (violet); and Concentration 20ppb (red).*

absorbance measurements for the four standard solutions of Cd (see Figure 3.4). The repeated measures for the peak absorbance were generated given a multivariate normal distribution $\boldsymbol{\beta}_t$ with mean vector

$$\boldsymbol{\mu} = \begin{bmatrix} 0.72 \\ 16.448 \\ -0.288 \end{bmatrix}$$

and variance-covariance matrix

$$\boldsymbol{\Sigma} = \sigma^2 \begin{bmatrix} 0.17966 & -0.03435 & 0.00131 \\ -0.03435 & 0.01473 & -0.00069 \\ 0.00131 & -0.00069 & 0.00003 \end{bmatrix},$$

where $\sigma^2 = 4.7$, the residual variance from the ordinary least squares fit of the original data. The time series of the posterior estimates for Cd is given in Figure 3.5. The square root of the mean squared error is 0.0025 with an average 95% credible interval [9.8 10.2] which is

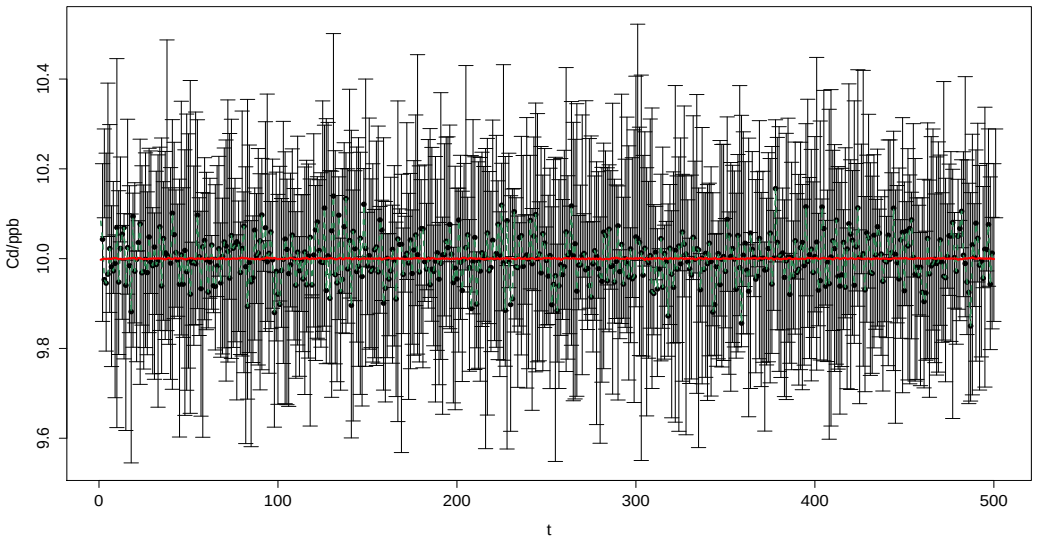


Figure 3.5: *Time series of Cd estimates (black dots connected by green lines) with corresponding 95% credible interval at time t. The "true" Cd of 10 ppb is given by the red line.*

shorter than the 95% confidence interval of [9.7 10.3] reported by Lundberg and de Maré (1980).

3.4.2 Example 2

We turn to an example in microwave radiometry to also demonstrate the usefulness of the proposed dynamic calibration approach. Engineers and scientist commonly use microwave radiometers to measure electromagnetic radiation. This radiant power is emitted by some source or a particular surface such as ice or land surface. Radiometers are very sensitive instruments that are capable of measuring extremely low levels of radiation. The transmission source of the radiant power is the target of the radiometers antenna. When a scene, such as terrain, is observed by a microwave radiometer, the radiation received by the antenna is partly due to self-emission by the scene and partly due to the reflected radiation originating from the surroundings (Ulaby et al. 1981). This source may be cosmic background radiation, ocean surface, or a heated surface used for the purpose of calibration.

Calibration is required due to the fact that the current electronic hardware is unable to maintain a stable input/output relationship. For space observing instruments, stable calibration without any drifts is a key to detect proper trends of climate (Imaoka *et al.* 2010). Problems such as amplifier gain instability and exterior temperature variations of critical components may cause this relationship to drift over time (Bremer 1979). During the calibration process, the radiometer receiver measures the voltage output power $v(t)$, and

its corresponding input temperature of a known reference. Two or more known reference temperatures are needed for calibration of a radiometer. Ulaby *et al.* (1981), and Racette and Lang (2005) state that the relationship between the output, $v(t)$, and the input, T_A, is approximately linear, and can be expressed as

$$\hat{T}_A = \hat{\beta}_0 + \hat{\beta}_1 v(t)$$

where $\hat{\beta}_0$ and $\hat{\beta}_1$ are the least square estimates for the regression parameters, $\hat{T}_A$ is the estimated value of the brightness temperature, and $v(t)$ is the observed output voltage. Using this relationship, the output value, $v(t)$, is used to derive an estimate for the input, T_A (Racette and Lang, 2005).

It is of interest to develop a calibration approach that can detect gain abnormalities, and/or correct for slow drifts that affect the quality of the instrument measurements. To demonstrate the dynamic approach in terms of application appeal, Rivers and Boone (2014) used the dynamic approach to characterize a calibration target over time for a microwave radiometer. The data used for this example was collected during a calibration experiment that was conducted on the Millimeter-wave Imaging Radiometer (MIR) (Racette et al. 1995). The purpose of the experiment was to validate predictions of radiometer calibration.

We extend the work of Rivers and Boone (2014) to the case when the relationship

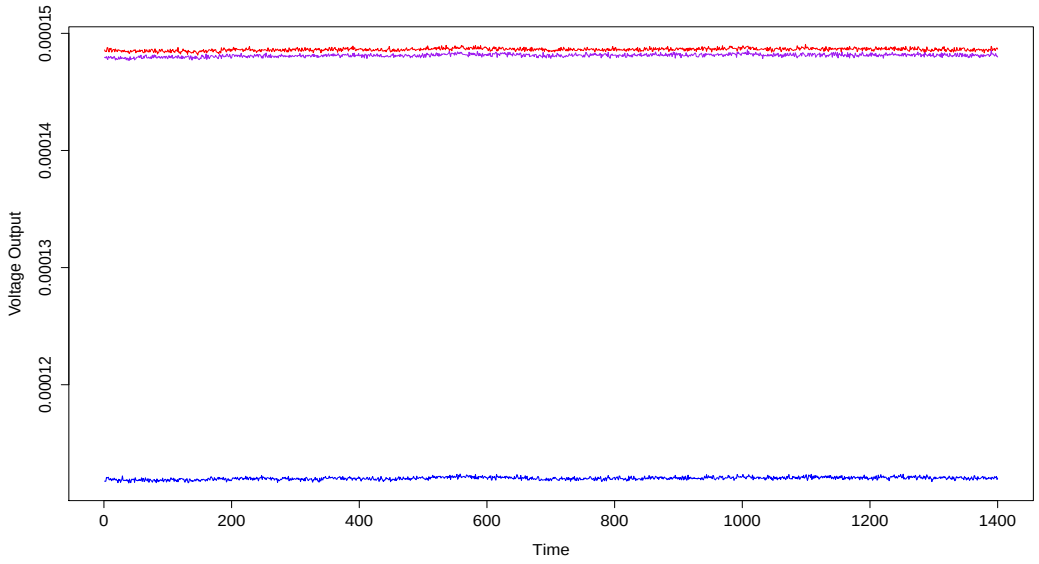


Figure 3.6: *Power output measurements: V_{cryo} (blue); V_{amb} (violet); and V_{warm} (red).*

between the reference measurment and the response is nonlinear by examining a data set similar to one created for a laboratory based calibration experiment. The National Institute of Standards and Technology (N.I.S.T.) conducted a bench calibration experiment designed

at studying calibration methods, but due to intellectual property rights the data is not free to use, so with the aid of computers we simulated data with the attributes of the N.I.S.T. data. The data set consists of three temperature references and the corresponding three power output measurements collected over 1400 time periods. The references temperatures are as follows:

- $\mathrm{T}_{cryo} = 84.3^{\circ}K$
- $\mathrm{T}_{amb} = 296.2^{\circ}K$
- $\mathrm{T}_{warm} = 300.7^{\circ}K$

with summary statistics for the corresponding output measurements as

	V_{cryo}	V_{amb}	V_{warm}
$\bar{v}$	0.0001120096	0.0001481137	0.0001486190
σ_v	0.0000001280	0.0000001308	0.0000001236

The time series plots for the observed output measurements are provided in Figure 3.6.

In Figure 3.7 we plot the output measurements against the reference temperatures and show the calibration curve by the dotted line. One may suggest that a linear fit is appropriate but the experiment that inspired this data was created such that the quadratic term β_{2t} would be significant.

We demonstrate the dynamic calibration method on the simulated radiometer data by setting $y_{0t} = 0.0001347169$ which corresponds to $x_{0t} = 200^{\circ}K$ given a stable system without any drift in the dynamic regression parameters. An assessment of the dynamic method is conducted at the end of the time series when $t = 1400$. After employing Algorithm 4 on the data we get a mean value for $\hat{x}_{0t}$ across time as 199.5695 with a standard deviation $\sigma_{x_{0t}} = 1.53135$. It is of interest to know if the addition of references would improve the estimation process because the placement of the references in the experiment conducted at NIST may not completely capture the degree of the nonlinearity. In Figure 3.8 we added two additional reference measurements. We place the additional measurements as follow:

- $\mathrm{T}_{135} = 135^{\circ}K$
- $\mathrm{T}_{245} = 245^{\circ}K$

with the corresponding output measures listed in the table below

	V_{135}	V_{245}
$\bar{v}$	0.0001228344	0.0001415994
σ_v	0.0000001257	0.0000001233

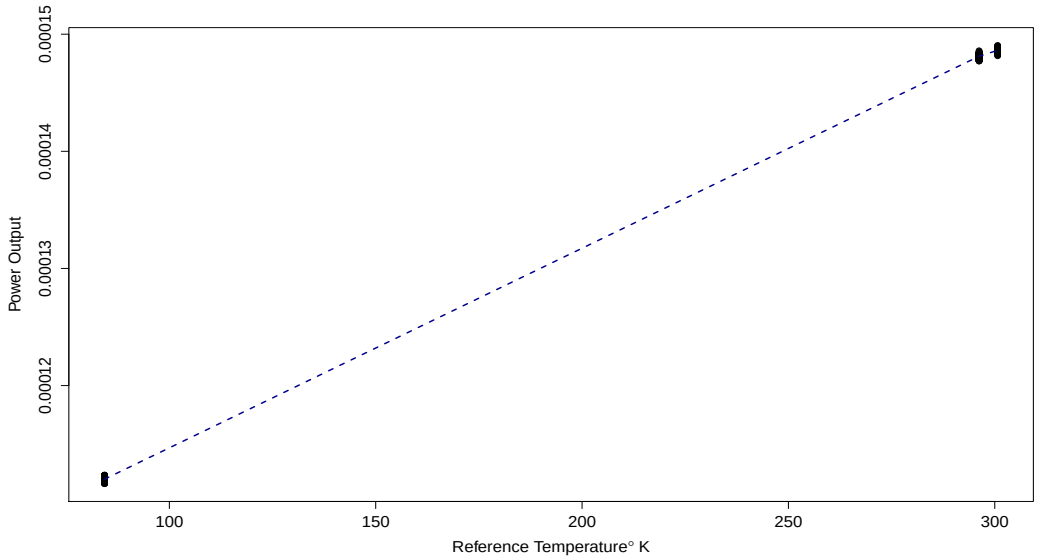


Figure 3.7: *3-Point Calibration model used to detect nonlinearity.*

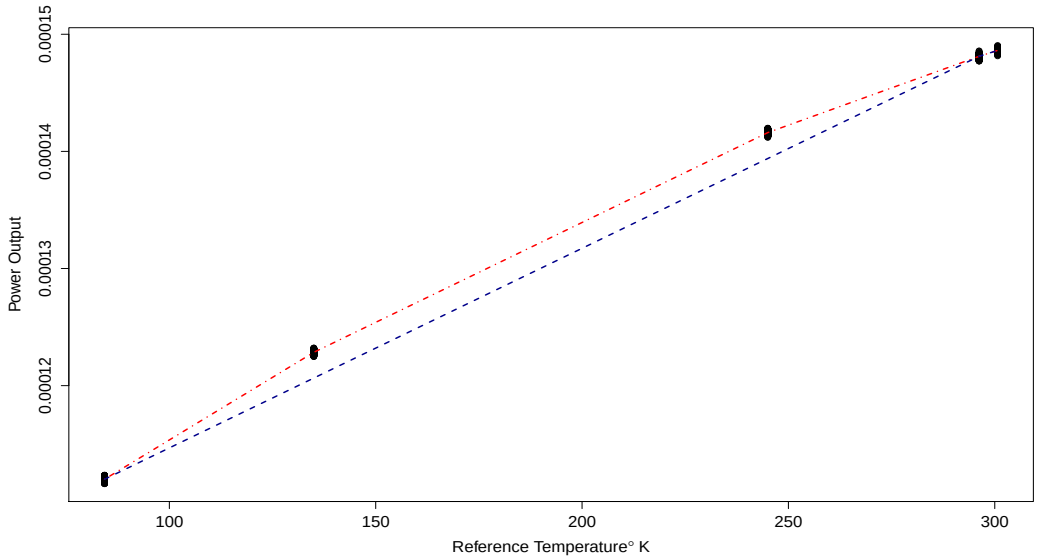


Figure 3.8: *5-Point Calibration model versus 3-Point Calibration model.*

The dynamic calibration method was conducted on the data set with the inclusion of the additional reference measurements yielding a mean value across time of $\hat{x}_{0t} = 200.1122$ which is closer to the assumed true temperature measure of $200°K$ and a standard deviation of $\sigma_{x_{0t}} = 0.4042387$. In Table 3.5 we compare the performance criterions from Section 3.3 side-

by-side for the data when using a 3-point calibration model and 5-point calibration model.

In Table 3.5 it is easy to see that the inclusion of the two reference measurements

Table 3.5: *Summary of the 3-point calibration model versus the 5-point calibration model.*

Model	MSE	IW	CP
3-Point	2.5130986	2.2480498	0.4450423
5-Point	0.1602064	1.8286866	0.9699140

greatly improved the estimation of the calibration distributions across time. This is evident by the decrease in the mean square error. The interval width is shorter for the 5-point model because, as stated before, the standard deviation when the 2 measurements were added is significantly smaller than when not including them. The greatest improvement took place in the coverage probability (CP) because it nearly doubled by pacing the references between the endpoint measurements.

The results shown in this section indicate that dynamic calibration can be used to determine concentrations, reference temperatures, or any other unknown measurement source where time period of calibration is significant. Until Rivers and Boone (2014) introduced the dynamic method, statistical calibration has been considered from a static point of view. To add to the discussion of the dynamic Bayesian nonlinear calibration, next we consider a case where concern may arise when the observed measurement y_{0t} approaches the vertex. This sort of issue is considered to be problematic, thus we are concerned with how the dynamic method performs in this scenario.

3.5 Future works and other considerations

Before concluding this work we would be remiss not to consider the case when the observed measurement y_{0t} approaches the vertex. We would like to understand how the calibration method will perform under such a condition. To understand the behavior of the vertex in a dynamic sense we look at the time-varying quadratic equation in vertex form,

$$y_t = \hat{a}_t(x - \hat{h}_t)^2 + \hat{k}_t \tag{3.19}$$

where

$$\begin{aligned} \hat{a}_t &= \hat{\beta}_{2t}, \\ \hat{h}_t &= \frac{-\hat{\beta}_{1t}}{2\hat{\beta}_{2t}}, \\ \hat{k}_t &= \hat{\beta}_{0t} - \frac{-\hat{\beta}_{1t}^2}{4\hat{\beta}_{2t}}, \end{aligned}$$

and the time-varying vertex is $(\hat{h}_t,\ \hat{k}_t)$.

The estimated dynamic regression parameters are used to derive the joint distribution

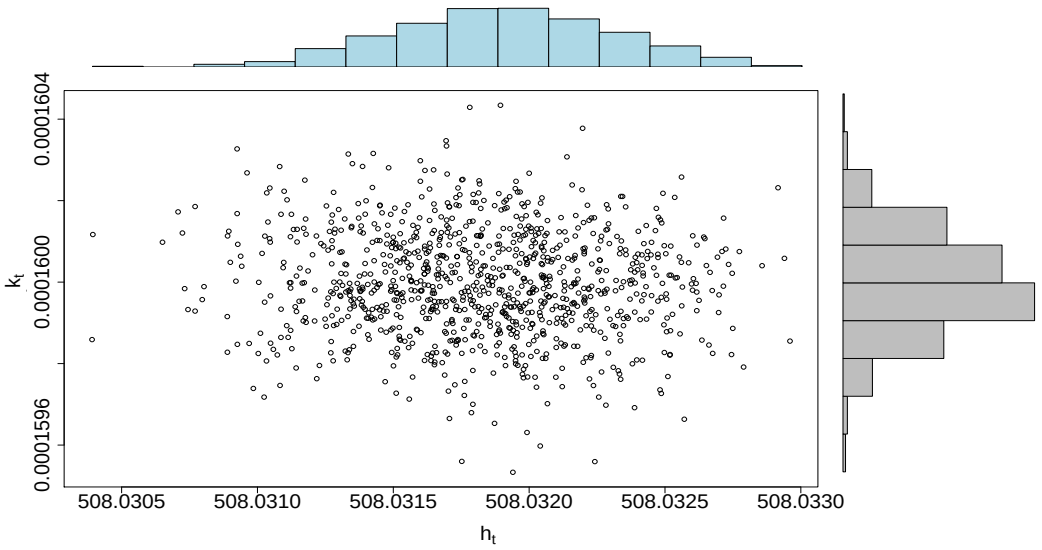


Figure 3.9: *Joint distribution of* $(\hat{h}_t,\ \hat{k}_t)$ *with marginal distributions of* $\hat{h}_t$ *and* $\hat{k}_t$.

of $(\hat{h}_t,\ \hat{k}_t)$ given in Figure 3.9. So, we extend the simulated radiometer data previously used in Section 3.4 by placing a reference measurement near the vertex at $508°K$. A series of observed power outputs y_{0t} is generated from the following quadratic equation:

$$y_{0t} = \hat{\beta}_{0t} + \hat{\beta}_{1t}(508) + \hat{\beta}_{2t}(508)^2, \tag{3.20}$$

where $\hat{\beta}_{0t}$, $\hat{\beta}_{1t}$ and $\hat{\beta}_{2t}$ are the estimates of the time dependent regression parameters. In Figure 3.10 the black horizontal line represents the possible maximum power output V_t given the vertex is $(\hat{h}_t,\ \hat{k}_t)$ at time t.

We used the dynamic calibration method described by Algorithm 4 to examine just how the method would perform at estimating the temperature value of $508°K$. In Figure 3.11 we see that the method performs poorly from a mean squared error and coverage probability point of view. The method is restricted from deriving distributions that violate the quadratic behavior by not calculating credible intervals that go beyond any real possible value, meaning that the upper credible limit x_{0t}^U will be less than the true value for x_{0t} with $P(x_{0t}^U < x_{0t}) = 1$. Clearly, it is not advisable to perform a statistical calibration experiment so close to the vertex. Calibration, whether static or dynamic is best conducted when the calibration target is near the center of the references domain space.

In this paper, we have proposed a robust nonlinear calibration algorithm to approximate the posterior calibration distribution of an unknown reference measurement. Our algorithm

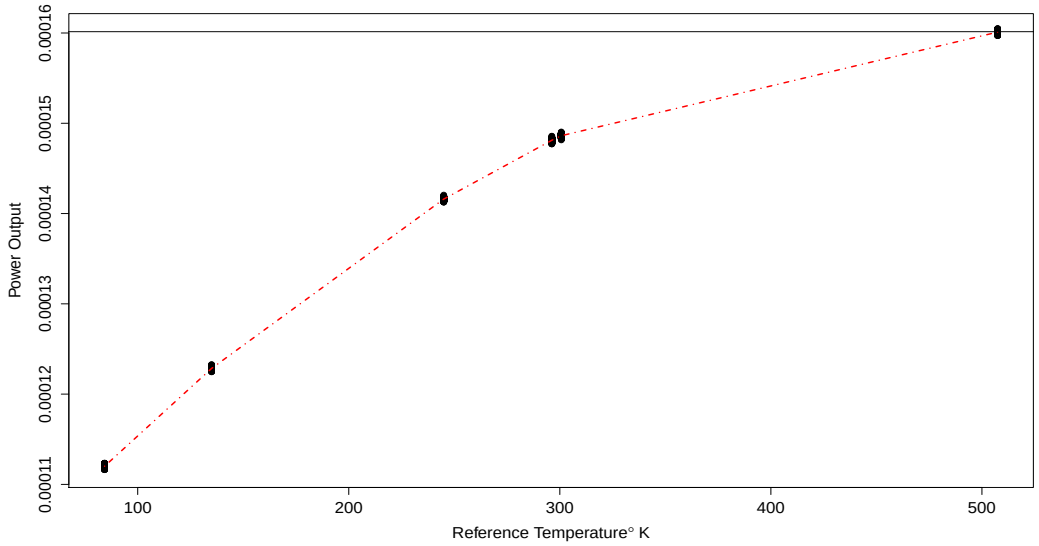


Figure 3.10: *Quadratic calibration model with observation y_{0t} near vertex.*

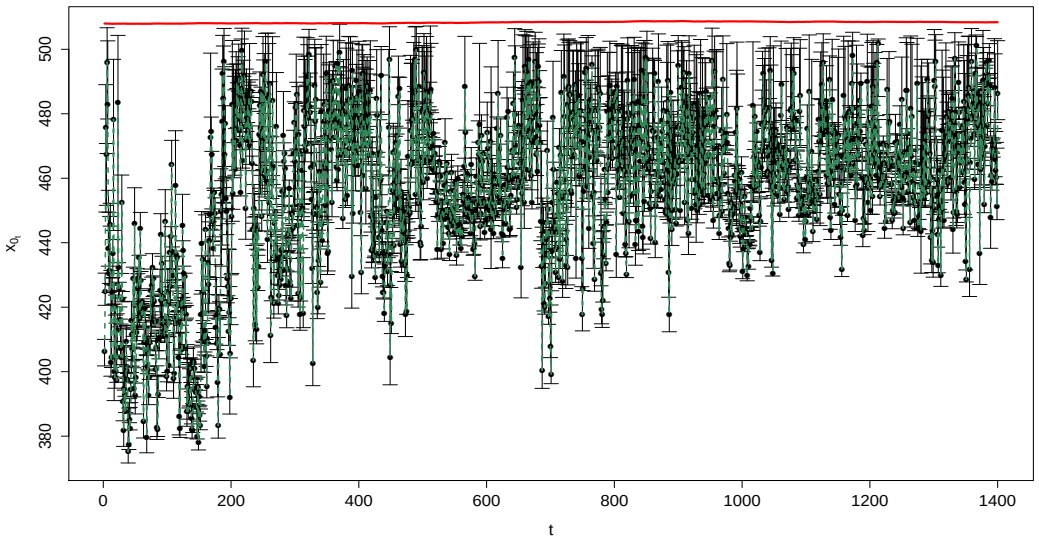


Figure 3.11: *Dynamic results of quadratic calibration model with observation y_{0t} near vertex.*

has the advantage of being adaptable to various calibration experiments. Our approach successfully combines Bayesian time series analysis, nonlinear calibration, and sampling theory. Indeed, we showed that the statistical calibration problem in the presense of a nonlinear relationship can be thought of as a time series problem and posterior estimates can be derived

over time.

3.6 References

[1] Aitchison, J., and Dunsmore, I. R. (1975). *Statistical prediction analysis*. Cambridge University Press.

[2] Albert, J. (2007). *Bayesian Computation with R* (Vol. 747389981). New York: Springer.

[3] Berkson, J. (1969). Estimation of a Linear Function for a Calibration Line; Consideration of a Recent Proposal. *Technometrics*. **11**, 649-660.

[4] Bernardo, J. M. (1979). Reference posterior distributions for Bayesian inference. *Journal of the Royal Statistical Society. Series B (Methodological)*. 113-147.

[5] Bremer, J. C. (1979). Improvement of Scanning Radiometer Performance by Digital Reference Averaging. *Instrumentation and Measurement, IEEE Transactions on* 28.1: 46-54.

[6] Brown, P. J. (1982). Multivariate calibration. *Journal of the Royal Statistical Society B*. **44**, 287-321.

[7] Cervenka, P.O., and Massa, L. (1994). Research and Development Report: *Nonlinear Calibration of an Infrared Radiometer*. Paper presented at Workshop on Infrared Ship Signature Modeling (IRSSM), organized by the NATO Research Study Group on Maritime Infrared Target and Background Signatures, Measurements and Characteristics (RSG.5), Paris, France.

[8] du Plessis, J. L., and van der Merwe, A. J. (1996). Bayesian calibration in the estimation of the age of rhinoceros. *Annals of the Institute of Statistical Mathematics*. **48**(1), 17-28.

[9] Eisenhart, C. (1939). The interpretation of certain regression methods and their use in biological and industrial research. *Annals of Mathematical Statistics*. **10**, 162-186.

[10] Eno, D. R. (1999). Noninformative Prior Bayesian Analysis for Statistical Calibration Problems. Doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

[11] Eno, D. R., and Ye, K. (2000). Bayesian reference prior analysis for polynomial calibration models. *Test*. **9**(1), 191-208.

[12] Fong, Y., Wakefield, J., De Rosa, S., and Frahm, N. (2012). A Robust Bayesian Random Effects Model for Nonlinear Calibration Problems. *Biometrics*. **68**(4), 1103-1112.

[13] François, N., Govaerts, B., and Boulanger, B. (2004). Optimal designs for inverse prediction in univariate nonlinear calibration models. *Chemometrics and intelligent laboratory systems*, **74**(2), 283-292.

[14] Givens, G. H., and Hoeting, J. A. (2005). *Computational Statistics*. John Wiley and Sons.

[15] Halperin, M. (1970). On Inverse Estimation in Linear Regression. *Technometrics*. **12**, 727-736.

[16] Hibbert, D. B. (2006). The uncertainty of a result from a linear calibration. *Analyst*, **131**(12), 1273-1278.

[17] Hoadley, B. (1970). A Bayesian look at Inverse Linear Regression. *Journal of the American Statistical Association*. **65**, 356-369.

[18] Hunter, W., and Lamboy, W. (1981). A Bayesian Analysis of the Linear Calibration Problem. *Technometrics*. **23**, 323-328.

[19] Imaoka, K., Kachi, M., Kasahara, M., Ito, N., Nakagawa, K., and Oki, T. (2010). Instrument Performance and Calibration of AMSR-E and AMSR2. *International Archives of the Photogrammetry, Remote Sensing and Special Information Science*. **38**(Part 8).

[20] Kirkup, L., and Mulholland, M. (2004). Comparison of linear and nonlinear equations for univariate calibration. *Journal of Chromatography A*. **1029**(1), 1-11.

[21] Krutchkoff, R. G. (1967). Classical and Inverse Regression Methods of Calibration. *Technometrics*. **9**, 425-439.

[22] Lavagnini, I., and Magno, F. (2007). A statistical overview on univariate calibration, inverse regression, and detection limits: Application to gas chromatography/mass spectrometry technique. *Mass Spectrometry Reviews*. **26**(1), 1-18.

[23] Lim, J. H., and Yun, H. I. (2010). Comparison of linear and nonlinear equation for the calibration of roxithromycin analysis using liquid chromatography/mass spectrometry. Korean Journal of Veterinary Research. **50**(1), 11-17.

[24] Lundberg, E., and de Maré, J. (1980). Interval estimates in the spectroscopy calibration problem. *Scandinavian Journal of Statistics*. **7**(1) 40-42.

[25] Lwin, T. and Maritz, J. S. (1982). An Analysis of the Linear Calibration Controversy from the Perspective of Compound Estimation. *Technometrics*. **24**, 235-242.

[26] Martinelle, S. (1970). On the Choice of Regression in Linear Calibration. *Technometrics*. **12**(1), 157-161.

[27] Merkle, W. (1983). Statistical methods in regression and calibration analysis of chromosome aberration data. *Radiation and environmental biophysics*. **21**(3), 217-233.

[28] Naylor, J. C., and Smith, A. F. (1982). Applications of a method for the efficient computation of posterior distributions. *Applied Statistics*. 214-225.

[29] Osborne, C. (1991). Statistical Calibration: A Review. *International Statistical Review*. **59**(3), 309-336.

[30] Racine-Poon, A. (1988). A Bayesian approach to nonlinear calibration problems. *Journal of the American Statistical Association*. **83**(403), 650-656.

[31] Rivers, D. L., and Boone, E. L. (2014). A Dynamic Approach to Linear Statistical Calibration with an Application in Microwave Radiometry. *arXiv preprint arXiv:1406.7863*. Submitted to *Journal of Statistical Computation and Simulation*.

[32] Racette, P., and Lang, R. H. (2005). Radiometer Design Analysis Based Upon Measurement Uncertainty. *Radio Science*. **40**(5).

[33] Rubin, D. B. (1987). The calculation of posterior distributions by data augmentation: Comment: A noniterative sampling/importance resampling alternative to the data augmentation algorithm for creating a few imputations when fractions of missing information are modest: The SIR algorithm. *Journal of the American Statistical Association*, 543-546.

[34] Smith, R. L., and Corbett, M. (1987). Measuring marathon courses: an application of statistical calibration theory. *Applied Statistics*. 283-295.

[35] Smith, A. F., and Gelfand, A. E. (1992). Bayesian statistics without tears: a sampling-resampling perspective. *The American Statistician*, **46**(2), 84-88.

[36] Tierney, L., and Kadane, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association*. **81**(393), 82-86.

[37] ter Braak, C. J. (1995). Nonlinear methods for multivariate statistical calibration and their use in palaeoecology: a comparison of inverse (k-nearest neighbours, partial least squares and weighted averaging partial least squares) and classical approaches. *Chemometrics and Intelligent Laboratory Systems*, **28**(1), 165-180.

[38] Ulaby, F. T., Moore, R. K., and Fung, A. K. (1981). *Microwave Remote Sensing:Active and Passive, Vol. I: Fundamentals and Radiometry*. Addison-Wesley, Advanced Book Program, Reading, Massachusetts.

[39] Weinreb, M. P., Hamilton, G., Brown, S., and Koczor, R. J. (1990). Nonlinearity corrections in calibration of Advanced Very High Resolution Radiometer infrared channels. *Journal of Geophysical Research: Oceans (19782012)*. **95**(C5), 7381-7388.

[40] West, M., Harrison, P. J., and Migon, H. S. (1985). Dynamic generalized linear models and Bayesian forecasting. *Journal of the American Statistical Association*, **80**(389), 73-83.

[41] West, M. and Harrison, P.J. (1997). Bayesian Forecasting and Dynamic Linear Models, 2nd Edition. Springer-Verlag, New York.

[42] Williams, E. J. (1969). A Note on Regression Methods in Calibration. *Technometrics*. **11**(1), 189-192.

[43] Ziemer, R. R., and Strauss, D. (1978). A statistical approach to instrument calibration. USDA Forestry Service Technical Report.

Chapter 4

Dynamic Bayesian Statistical Analysis for Multi-Univariate Calibration

Abstract

In this paper the dynamic approach to the linear statistical calibration problem is extended from univariate to a special case of the multivariate calibration problem. The focus is on the *multi-univariate linear calibration*, in which the response variable is multivariate and the explanatory variable is univariate. The properties of the dynamic multi-univariate calibration model are compared to static frequentist and Bayesian multivariate approaches through a Monte Carlo simulation study.

4.1 Introduction

Just as simple linear regression can be extended to multiple regression and multivariate regression, the linear calibration problem can be extended to more general settings. The *multivariate linear calibration problem* is similar to the univariate calibration problem, but in the multivariate case it is assumed that q response variables depend linearly on p regressor variables. As before, the data for the calibration experiment are collected in two stages. At the first stage, n observations $(\mathbf{X}_i, \mathbf{Y}_i)$ are made. Here, each row of $\mathbf{X}_i$ is a $p \times 1$ vector, and each $\mathbf{Y}_i$ is a $q \times 1$ vector. It is assumed that each element of $\mathbf{X}_i$ is precisely determined, while the elements of $\mathbf{Y}_i$ are measurements which are subject to error. At the second stage of data collection, one or more observations of $\mathbf{Y}$ are made, and interest centers on the determination of the unknown $\mathbf{x}_0'$ vector giving rise to these observations. A new concern comes to light as we extend the univariate calibration problem to the multivariate case, concerning the relative sizes of q and p. Specifically, if the unknown value $\mathbf{x}_0'$ is to be uniquely determined, we require that the number of responses q be at least as large as the number of regressors p.

When it comes to on-line measurements, multivariate calibration with instruments employing the framework described above may be compromised by instrumental, process and operational drifts that are not seen during off-line calibration. This may render the calibration model unsuitable for prediction of key components (Gujral et al. 2009). Calibrations are never concluded once and for all. Instrument characteristics are altered by time and use and calibration must be viewed as an iterative process as long as the instrument is in use (Cervenka and Massa, 1994). One solution to this problem is to use a multivariate calibration approach that is dynamic by way of its ability to correctly predict an unknown reference value $\mathbf{x}_0'$ in the presence of systematic variations that may be due to the effect of temperature, pressure or aging of the instrument.

The paper is organized as follows. The basic principles of multivariate calibration from both the frequentist and Bayesian perspectives are reviewed in Section 4.2. Section 4.3 presents the development of the dynamic multi-univariate approach to the statistical calibration problem. In Section 4.5 the results are presented from a simulation study where the dynamic multi-univariate method is evaluated along with static multivariate approaches. A conclusion is given in Section 4.6.

4.2 Multivariate Calibration Model

Multivariate calibration was studied extensively by Brown (1982). In his paper, procedures for multivariate calibration that are comparable to the *classical* and *inverse* procedures for univariate calibration were given. Here the term multivariate is intended to mean multivariate in both the X's and the Y's. Suppose there are n observations in the *controlled* calibration experiment, q response variables $Y_1, Y_2, \ldots, Y_q$ and p explanatory variables $X_1, X_2, \ldots, X_p$, and $q \geq p$. Then a linear model may be expressed as

$$\mathbf{Y} = \mathbf{1}\boldsymbol{\alpha}' + \mathbf{X}\mathbf{B} + \mathbf{E}, \tag{4.1}$$

where $\mathbf{Y}$ and $\mathbf{E}$ are $(n \times q)$ random matrices, $\mathbf{X}$ is a $n \times p$ matrix of fixed constant, $\mathbf{1}$ is a $(n \times 1)$ vector of ones, $\mathbf{B}$ is a $p \times q$ matrix of unknown parameters, and $\boldsymbol{\alpha}$ is a $q \times 1$ vector of unknown parameters.

Brown (1982) assumes that

$$\sum_{i=1}^{n} x_{ij} = 0, \qquad \frac{1}{n}\sum_{i=1}^{n} x_{ij}^2 = 1, \qquad j = 1, \ldots, p, \tag{4.2}$$

which means that the columns of $\mathbf{X}$ are standardized and have average sum of squares equal one. The classical estimator $\hat{\mathbf{B}}$ is given by

$$\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \quad \text{and} \quad \hat{\boldsymbol{\alpha}} = \bar{\mathbf{y}}.$$

The model for the prediction experiment is given by

$$\mathbf{y}_0 = \mathbf{1}\boldsymbol{\alpha}' + \mathbf{1}\mathbf{x}_0'\mathbf{B} + \mathbf{E}^{\star}, \tag{4.3}$$

where $\mathbf{y}_0$ and $\mathbf{E}^\star$ are $(m \times q)$ random matrices, $\mathbf{x}_0$ is a $p \times 1$ vector of unknown values and $\mathbf{1}$ is a $(m \times 1)$ vector of ones. It is wished to draw inferences about $\mathbf{x}_0$. If $\mathbf{e}_i$ is the i^{th} row of $\mathbf{E}$, it is assumed that

$$E(\mathbf{e}_i) = 0, \;\; E(\mathbf{e}_i \mathbf{e}_i^T) = \mathbf{\Gamma} \text{ and } \mathbf{e}_i \sim N(\mathbf{0}, \mathbf{\Gamma}), \qquad i = 1, 2, \ldots, n. \tag{4.4}$$

If $\mathbf{e}_j$ is the j^{th} row of $\mathbf{E}^\star$, the $\mathbf{e}_j^\star$ satisfy the above also and it is assumed that they are independent of the $\mathbf{e}_i$ (Brown 1982; Özyurt and Erar 2003).

It was also shown by using the *classical* approach similar to Eisenhart (1939) that a $100(1-\gamma)\%$ confidence region for $\mathbf{x}_0$ is

$$(\mathbf{y}_0 - \hat{\boldsymbol{\alpha}} - \hat{\mathbf{B}}'\mathbf{x}_0)' \mathbf{S}^{-1} (\mathbf{y}_0 - \hat{\boldsymbol{\alpha}} - \hat{\mathbf{B}}'\mathbf{x}_0) / \sigma^2(\mathbf{x}_0) \leq \frac{q}{v} F_{\gamma,q,v},$$

where $\mathbf{S}$ is a $(q \times q)$ matrix given by

$$\mathbf{S} = \hat{\mathbf{E}}'\hat{\mathbf{E}} = (\mathbf{Y} - \mathbf{1}\hat{\boldsymbol{\alpha}}' - \mathbf{X}\hat{\mathbf{B}})'(\mathbf{Y} - \mathbf{1}\hat{\boldsymbol{\alpha}}' - \mathbf{X}\hat{\mathbf{B}}) \tag{4.5}$$

with $v = n - p - q$ degrees of freedom and

$$\sigma^2(\mathbf{x}_0) = \frac{1}{m} + \frac{1}{n} + \mathbf{x}_0'(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_0$$

and $F_{\gamma,q,v}$ is the upper $100(1-\gamma)\%$ point of the standard F distribution on q and v degrees of freedom.

Thus the estimation of the unknown vector $\mathbf{x}_0$ is obtained in the form

$$\hat{\mathbf{x}}_0 = (\hat{\mathbf{B}}\mathbf{S}^{-1}\hat{\mathbf{B}}')^{-1}\hat{\mathbf{B}}\mathbf{S}^{-1}(\mathbf{y}_0 - \bar{\mathbf{y}}). \tag{4.6}$$

Brown (1982) also obtains an estimation of $\mathbf{x}_0$ using an inverse estimation method like that of Krutchkoff (1967) where the model is

$$\hat{\mathbf{X}} = \mathbf{Y}\hat{\mathbf{B}}_k$$

and $(\mathbf{X} - \hat{\mathbf{X}})'(\mathbf{X} - \hat{\mathbf{X}})$ is minimized using the least squares method. The least squares estimate of $\mathbf{B}_k$ is

$$\mathbf{B}_k = (\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{Y}'\mathbf{X}.$$

For a given $1 \times p$ dimensioned observation $\mathbf{y}_0$, $\hat{\mathbf{x}}_{0,k}$ is given by

$$\hat{\mathbf{x}}_{0,k} = \mathbf{X}'\mathbf{Y}(\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{y}_0'. \tag{4.7}$$

In addition, Brown (1982) provided a multivariate calibration extension to the Bayesian approach to the calibration problem by Hoadley (1970). In this Bayesian solution, Brown (1982) used a noninformative prior for all model parameters except the unknown $\mathbf{x}_0$,

$$\pi(\mathbf{B}, \boldsymbol{\alpha}, \mathbf{\Gamma}, \mathbf{X}_0) = \pi(\mathbf{B}, \boldsymbol{\alpha}, \mathbf{\Gamma})\pi(\mathbf{X}_0), \tag{4.8}$$

where a Jeffrey's invariant prior (Jeffreys, 1946) is assumed,

$$\pi(\mathbf{B}, \boldsymbol{\alpha}, \boldsymbol{\Gamma}) \propto |\boldsymbol{\Gamma}|^{-(q+1)} \tag{4.9}$$

which is the multivariate equivalent to the noninformative prior for $(\alpha, \beta_0, \sigma^2)$ (i.e., $\pi(\alpha, \beta_0, \sigma^2) \propto \sigma^{-2}$ proposed by Hoadley (1970)). Defining the model by Equations (4.1), (4.3), (4.2), (4.4), and (4.8), and by assuming

$$\pi(\mathbf{x}_0|\mathbf{X}) = \pi(\mathbf{x}_0),$$

(the $\mathbf{X}$ values provide no information on $\mathbf{x}_0$), Brown (1993) derives the conditional posterior for $\mathbf{x}_0$ as

$$\pi(\mathbf{x}_0|\mathbf{Y}', \mathbf{Y}, \mathbf{X}) \propto \pi(\mathbf{x}_0)L(\mathbf{x}_0),$$

where $L(\mathbf{x}_0)$ is the predictive distribution of $\bar{\mathbf{y}}_0'$. Assuming the standard natural conjugate prior in Equation (4.9), the predictive distribution is

$$L(\mathbf{x}_0) = \frac{\sigma^2(\mathbf{x}_0)^{\nu/2}}{\left[\sigma^2(\mathbf{x}_0) + (\bar{\mathbf{y}}_0' - \hat{\mathbf{B}}'\mathbf{x}_0)'\mathbf{S}^{-1}(\bar{\mathbf{y}}_0' - \hat{\mathbf{B}}'\mathbf{x}_0)\right]^{(\nu+q)/2}}, \tag{4.10}$$

where $\mathbf{Y}$ is centered in the calibration experiment.

Brown (1982) studied a specific multivariate Student's t distribution as the prior for $\mathbf{x}_0$ because it knocks out the numerator of Equation (4.10) and gives a Student posterior distribution for $\mathbf{x}_0$. Brown (1982) extends results from Hoadley (1970) in the following theorem:

Theorem 1. *Suppose a priori* $\mathbf{x}_0$ *has a multivariate Student distribution,* $T_{\nu-p}[\mathbf{0}, (1 + 1/n)\mathbf{X}'\mathbf{X}]$*, then the posterior distribution of* $\mathbf{x}_0$ *is* $T_{\nu+q-p}[\hat{\mathbf{x}}_{0,k}, (1+1/n+\mathbf{y}_0'(\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{y}_0)(\mathbf{G} + \mathbf{K})^{-1}]$*, where* $\mathbf{G} = (\mathbf{X}'\mathbf{X})^{-1}$*,* $\mathbf{H} = \hat{\mathbf{B}}\mathbf{S}^{-1}\hat{\mathbf{B}}'$ *and the columns of* $\mathbf{X}$*,* $\mathbf{Y}$ *have been centered.*

As was the case in the earlier work of Hoadley (1970), this prior was studied because the resulting mode of the marginal posterior distribution of $\mathbf{x}_0$ is equal to the inverse estimator of $\mathbf{x}_0$ in Equation (4.7). Brown did not propose any noninformative prior distributions for $\mathbf{x}_0$.

Several papers followed Brown (1982) that noted that there were a wide variety of approaches to multivariate calibration. More recently Naes et al. (1986) compared the multiple linear regression (MLR), ridge regression (RR), principal component regression (PCR) and partial least squares regression (PLSR) approaches with particular reference to the calibration of near infra-red (NIR) instruments. Brown and Sundberg (1987) considered an approach to multivariate calibration which involves the profile or maximum relative likelihood (Kalbfleisch and Sprott, 1970). An extension to Brown's (1982) multivariate calibration approach is proposed by Fox (1989). His work covers the situation where one is interested in calibrating for an unknown q-vector $\mathbf{X}$ on the basis of an observed p-vector $\mathbf{Y}$ given that $k \geq 1$ components are fixed in advance. Sundberg and Brown (1989) examined the case of

there being more variables than observations ($n < p + q + 1$) when assuming a standard multivariate linear regression model in equation 4.1. By way of an example, they used the NIR data of Fearn (1983).

du Plessis and van der Merwe (1995) developed a Bayesian approach to the multivariate and conditional calibration problems. They presented a multivariate Bayesian approach akin to that of the univariate calibration approach of Hunter and Lamboy (1981). For $\mathbf{\Gamma}$ unknown and $q > 1$, du Plessis and van der Merwe (1995) approximate the posterior distribution of $\mathbf{x}_0$ (defined in Equation (4.6)) through a simulation study because the exact marginal posterior distribution of $\mathbf{x}_0$ is difficult to derive and cannot be obtained in closed form. *A priori* they assume that little is known about the elements of $\mathbf{\Gamma}$. They further assume that the elements of $\mathbf{B}$ and those of $\mathbf{\Gamma}$ are independently distributed; that is

$$\pi(\mathbf{B}, \mathbf{\Gamma}) = \pi(\mathbf{B})\pi(\mathbf{\Gamma}),$$

where

$$\begin{aligned} \pi(\mathbf{B}) &\propto \text{constant}, \qquad \text{and} \\ \pi(\mathbf{\Gamma}) &\propto |\mathbf{\Gamma}|^{\frac{1}{2}(p+1)}. \end{aligned} \tag{4.11}$$

Using Equation (4.11), du Plessis and van der Merwe (1995) state that the marginal posterior distribution for $\mathbf{\Gamma}$ follows an inverted Wishart distribution,

$$\pi(\mathbf{\Gamma}|\mathbf{Y}) \propto |\mathbf{\Gamma}|^{\frac{\nu}{2}} e^{-\frac{1}{2}tr(\mathbf{\Gamma}^{-1})\mathbf{S}},$$

where $\mathbf{S}$ is defined in Equation (4.5) and $\nu = n - q + p$.

In the same spirit as Hunter and Lamboy (1981), du Plessis and van der Merwe (1995) write the posterior distribution as $\pi(\mathbf{y}_0' - \boldsymbol{\alpha}, \mathbf{B}, \mathbf{\Gamma}|\mathbf{Y}) = \pi(\mathbf{y}_0' - \boldsymbol{\alpha}|\mathbf{B}, \mathbf{\Gamma}, \mathbf{Y})\pi(\mathbf{B}|\mathbf{\Gamma}, \mathbf{Y})\pi(\mathbf{\Gamma}|\mathbf{Y})$, where

$$\begin{aligned} \mathbf{B}|\mathbf{\Gamma}, \mathbf{Y} &\sim N[\hat{\mathbf{B}}, \mathbf{C}_{11} \otimes \mathbf{\Gamma}] \\ (\mathbf{y}_0' - \boldsymbol{\alpha})|\mathbf{B}, \mathbf{\Gamma}, \mathbf{Y} &\sim N[\boldsymbol{\mu}^*, \mathbf{\Gamma}^*], \end{aligned}$$

where

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} c_{00}(1 \times 1) & \mathbf{c}_{01}(1 \times q) \\ \mathbf{c}_{10}(q \times 1) & \mathbf{C}_{11}(q \times q) \end{bmatrix},$$

and $\otimes$ is the Kronecker product. Through simulation, du Plessis and van der Merwe (1995) approximate the posterior distribution of $\mathbf{x}_0$ in the following way:

(i) Draw $\mathbf{\Gamma}^{-1}$ from a $\mathcal{W}^{-1}[\mathbf{S}, n - q - 1]$ distribution.

(ii) Given $\mathbf{\Gamma}$, draw $\mathbf{B}$ from a $N[\hat{\mathbf{B}}, \mathbf{C}_{11} \otimes \mathbf{\Gamma}]$ distribution.

(iii) Calculate the posterior density function

$$\pi(\mathbf{x}_0|\mathbf{B}, \mathbf{\Gamma}, \mathbf{Y}) = N[\boldsymbol{\mu}^*, \mathbf{\Gamma}^*], \tag{4.12}$$

where the mean vector

$$\boldsymbol{\mu}^* = (\hat{\mathbf{B}}\boldsymbol{\Gamma}^{-1}\hat{\mathbf{B}}')^{-1}\hat{\mathbf{B}}\boldsymbol{\Gamma}^{-1}\left[\left(\mathbf{y}_0 - \hat{\boldsymbol{\alpha}}\right) - \left(\mathbf{B} - \hat{\mathbf{B}}\right)'\mathbf{C}_{11}^{-1}\mathbf{c}_{10}\right] \tag{4.13}$$

and the covariance matrix

$$\boldsymbol{\Gamma}^* = \left(1 + \frac{1}{n}\right)(\hat{\mathbf{B}}\boldsymbol{\Gamma}^{-1}\hat{\mathbf{B}}')^{-1}. \tag{4.14}$$

(iv) Repeat steps (i) to (iii) m times.

By averaging over the m repetitions du Plessis and van der Merwe (1995) obtain an estimate of the marginal distribution of $\mathbf{x}_0$.

A multivariate Bayesian competitor to the conditional procedure derived by Fox (1989) is also given by du Plessis and van der Merwe (1995). Using the data of Brown (1982), Fox (1989), and du Plessis and van der Merwe (1995), a comparative analysis was done by Özyurt and Erar (2003) that studied how the conditional calibration technique of Fox (1989) and the classical and inverse multivariate techniques of Brown (1982) perform when outliers are present.

Considerable emphasis has been given to the multi-univariate case of the calibration problem where the response variable $\mathbf{Y}$ is multivariate and the regressor variable $\mathbf{X}$ is univariate. This multi-univariate model is denoted by

$$\mathbf{Y}_i \sim N_q(\boldsymbol{\alpha} + \boldsymbol{\beta}x_i, \boldsymbol{\Gamma}) \qquad i = 1, 2, \ldots, n.$$

Lieftinick-Koeijers (1988) encountered this form of the calibration problem in a biological study by Broekhuizen and Maaskamp (1979) where it was employed as a means of determining the unknown age of deceased hares given several measured criteria. Oman and Wax (1984) consider the multi-univariate calibration approach when estimating gestational age by ultrasound measurement of fetal bone lengths. A similar example is provided by Brown (1979). A generalization of the classical estimator was considered by Lieftinick-Koeijers (1988). Oman and Srivastava (1996) compare the exact mean square error (MSE) for both Brown's (1982) classical estimator and inverse estimator in multi-univariate linear calibration. Moments for the classical estimator are derived by Nishii and Krishnaiah (1988) for multi-univariate calibration. Srivastava (1995) considers a Bayesian approach to both the inverse and classical estimator in the multi-univariate calibration case. It is shown by Srivastava (1995) that the classical estimator is inadmissible whereas the inverse estimator is admissible. Takeuchi (1997) proposed a class of generalized inverse regression estimators in multi-univariate calibration. For the proposed class of estimators, Takeuchi (1997) derives the expressions of bias and mean square error (MSE).

All of these previously reviewed multivariate calibration methods are considered to be static estimators. A static estimator is one that is not designed to incorporate drifts or

systematic changes over time. With all of the methods mentioned here, a standard multivariate assumption is that $n \geq p+q+1$ to avoid singularity issues with the estimate for the residual covariance matrix $\hat{\mathbf{\Gamma}}$. Sundberg and Brown (1989) state that when the assumption $n \geq p+q+1$ is not satisfied, $\hat{\mathbf{\Gamma}}$ is singular, thus invalidating Equation (4.6). They further conclude that if $n \leq q$, Equation (4.7) will invalidated. Alternatively, Sundberg and Brown (1989) support the use of a g-inverse $\hat{\mathbf{\Gamma}}^{-}$ of $\hat{\mathbf{\Gamma}}$. It is warned that not all g-inverses will be allowed because many of the g-inverses of $\hat{\mathbf{\Gamma}}$ will be proper inverses of other covariance matrices, which could differ essentially from $\hat{\mathbf{\Gamma}}$ and make the multivariate method unstable. A dynamic consideration that does not suffer from singularity issues would allow calibration to be adaptive to changes in the system and iterative over time. In the next section the development of a dynamic multi-univariate calibration method is discussed.

4.3 Dynamic Multi-Univariate Calibration Model

Let $\{(\mathbf{X}, \mathbf{Y}_t)|t = 1, 2, \ldots, T\}$ be the reference measurements and responses in the multi-univariate calibration experiment at time t and suppose the relationship can be described by

$$\mathbf{Y}_t = (\mathbf{X} \otimes \mathbf{I}_q)\boldsymbol{\theta}_t + \boldsymbol{\epsilon}_t, \qquad \boldsymbol{\epsilon}_t \sim N[\mathbf{0}, \mathbf{E}] \tag{4.15}$$

where $\mathbf{Y}_t(rq \times 1) = \begin{bmatrix} y_{11t} \\ \vdots \\ y_{q1t} \\ y_{12t} \\ \vdots \\ y_{qrt} \end{bmatrix}$, $\boldsymbol{\theta}_t(dq \times 1) = \begin{bmatrix} \beta_{01t} \\ \vdots \\ \beta_{0qt} \\ \beta_{11t} \\ \vdots \\ \beta_{1qt} \end{bmatrix}$, $\boldsymbol{\epsilon}_t(rq \times 1) = \begin{bmatrix} \epsilon_{11t} \\ \vdots \\ \epsilon_{q1t} \\ \epsilon_{12t} \\ \vdots \\ \epsilon_{qrt} \end{bmatrix}$,

and $(\mathbf{X} \otimes \mathbf{I}_q)$ is the Kronecker product of a fixed $(r \times d)$ reference matrix $\mathbf{X}$ and a $q-$dimensional identity matrix $\mathbf{I}_q$. The elements of the $rq-$dimensional vector $\boldsymbol{\epsilon}_t$ are independently normally distributed error terms with mean $\mathbf{0}$ and $(rq \times rq)$ covariance matrix $\mathbf{E} = \sigma_E^2 \mathbf{I}_{rq}$. Equation (4.15) is known as the observation equation.

The dynamic regression parameter vector $\boldsymbol{\theta}_t = [\boldsymbol{\beta}_{0t}\ \mathbf{B}_{1t}]'$ captures the evolving relationship between $\mathbf{X}$ and $\mathbf{Y}_t$ over time. Brown (1993) shows that under normality $\boldsymbol{\beta}_{0t}$ and $\mathbf{B}_{1t}$ are independent,

$$\hat{\boldsymbol{\beta}}_{0t}\sqrt{r} - \boldsymbol{\beta}_{0t}\sqrt{r} \sim N_q[\mathbf{0}, \mathbf{\Gamma}] \tag{4.16}$$

and

$$\hat{\mathbf{B}}_t - \mathbf{B}_t \sim N[(\mathbf{X}'\mathbf{X})^{-1}, \mathbf{\Gamma}], \tag{4.17}$$

with $\mathbf{\Gamma}$ being defined in Equation (4.4). Therefore the evolution of the regression parameters is modeled as

$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \qquad t = 1, 2, \ldots, T, \tag{4.18}$$

where $\boldsymbol{\omega}_t$ is a $dq-$dimensional vector,

$$\boldsymbol{\omega}_t = \begin{bmatrix} \omega_{01t} \\ \vdots \\ \omega_{0qt} \\ \omega_{11t} \\ \vdots \\ \omega_{1qt} \end{bmatrix}.$$

The evolution error terms are independently normally distributed with mean $\mathbf{0}$ and variance-covariance matrix $\mathbf{W}$. The matrix $\mathbf{W}$ is block-diagonal

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_{\beta_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_{\beta_1} \end{bmatrix},$$

with $\sigma_W^2 \mathbf{I}_q = \mathbf{W}_{\beta_0} = \mathbf{W}_{\beta_1}$. Equation (4.18) is known as the system equation.

The observation vector $\mathbf{Y}_t$ and the dynamic regression parameter vector $\boldsymbol{\theta}_t$ are both random variables, thus the expected values $\hat{\mathbf{Y}}_t$ and $\hat{\boldsymbol{\theta}}_t$ are the means of their respective distributions and must be estimated sequentially. The one-step forecast for $\mathbf{Y}_t$ and posterior distributions $\boldsymbol{\theta}_t$ for each time t are as follows in Algorithm 5 (Dynamic Linear Models algorithm). See West et al. (1985), West and Harrison (1997), and Petris et al. (2009) for a more detailed discussion of Dynamic Linear Regression Models (DLRMs).

Furthermore, let $\{\mathbf{y}_{0t} | t = 1, 2, \ldots, T\}$ be the observation vector from the second stage of the calibration experiment corresponding to a single unknown reference value of interest x_{0t}. Akin to Brown's (1982) Equation (4.3) the second stage of experimentation is modeled as

$$\mathbf{y}'_{0t} = \boldsymbol{\beta}'_{0t} + x_{0t}\mathbf{B}'_t + \boldsymbol{\epsilon}'_{0t}, \qquad t = 1, 2, \ldots, T, \tag{4.19}$$

where $\mathbf{y}_{0t}$ is a $(q \times 1)$ vector of second-stage observations, $\boldsymbol{\beta}_{0t}$ is a $(q \times 1)$ vector of observation means from the first stage (e.g. $\boldsymbol{\beta}_{01} = \overline{\mathbf{Y}}_1$), $\mathbf{B}_t$ is $(q \times 1)$ vector of dynamic regression slope values. The terms of the error vector $\boldsymbol{\epsilon}_{0t}$ are assumed to be independently normally distributed with mean 0 and variance σ_E^2.

As with Hoadley (1970) and Brown (1982) it is assumed that the first stage of the calibration experiment is independent of the second stage, therefore x_{0t} is independent of $(\boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2)$ where the joint prior distribution is

$$\pi(x_{0t}, \boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2) = \pi(x_{0t})\pi(\boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2). \tag{4.20}$$

The posterior of $(x_{0t}, \boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2)$ is then given by

$$\begin{aligned} \pi(x_{0t}, \boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2 | \mathbf{y}_{0t}, \mathbf{Y}_t) &\propto f(\mathbf{y}_{0t}, \mathbf{Y}_t | x_{0t}, \boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2)\pi(x_{0t}, \boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2) \\ &\propto f(\mathbf{y}_{0t} | x_{0t}, \boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2)\pi(x_{0t}) f(\mathbf{Y}_t | \boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2)\pi(\boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2) \\ &\propto f(\mathbf{y}_{0t} | x_{0t}, \boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2)\pi(\boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2 | \mathbf{Y}_t)\pi(x_{0t}), \end{aligned} \tag{4.21}$$

Algorithm 5: Updating (*MDLRM*) Multivariate Dynamic Linear Regression Model

Initialize $t = 0$
{Initial information $(\boldsymbol{\theta}_0|D_0) \sim N_d[\mathbf{m_0}, \mathbf{C_0}]$}
Input: $\mathbf{m_0}$, $\mathbf{C_0}$, $\mathbf{E}$, $\mathbf{W}$
loop
  $t = t + 1$
  {Compute prior at t: $(\boldsymbol{\theta}_t|D_{t-1}) \sim N_d[\mathbf{a_t}, \mathbf{R_t}]$}
    $\mathbf{a}_t = \mathbf{m}_{t-1}$
    $\mathbf{R}_t = \mathbf{C}_{t-1} + \mathbf{W}$
  Input: $\mathbf{X}$
    $\mathbf{X} = \mathbf{X} \otimes \mathbf{I}_q$
  {Compute forecast at t: $(\mathbf{Y_t}|D_{t-1}) \sim N_r[\mathbf{f_t}, \mathbf{Q_t}]$}
    $\mathbf{f}_t = \mathbf{X}\mathbf{a}_t$
    $\mathbf{Q}_t = \mathbf{X}\mathbf{R}_t\mathbf{X}' + \mathbf{E}$
  Input: $\mathbf{Y}_t$
    $\boldsymbol{\beta}_{0t} = \overline{\mathbf{Y}}_t$
  {Compute forecast error $\mathbf{e}_t$}
    $\mathbf{e}_t = \mathbf{Y}_t - \boldsymbol{\beta}_{0t} - \mathbf{f}_t$
  {Compute adaptive gain matrix $\mathbf{A}_t$}
    $\mathbf{A}_t = \mathbf{Q}_t^{-1}\mathbf{X}\mathbf{R}_t$
  {Compute posterior at t: $(\boldsymbol{\theta}_t|D_t) \sim N_d[\mathbf{m}_t, \mathbf{C}_t]$}
    $\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_t\mathbf{e}_t$
    $\mathbf{C}_t = \mathbf{R}_t - \mathbf{A}_t'\mathbf{Q}_t\mathbf{A}_t$
  {Compute residual sum of products matrix $\mathbf{S}_t$ at t: $(\mathbf{S}_t|D_t) \sim \mathcal{W}[\mathbf{S}_t, n - q - 1]$}
    $\mathbf{S}_t = (\mathbf{Y}_t - \mathbf{1}\boldsymbol{\beta}_{0t}' - \mathbf{X}\mathbf{m}_t)'(\mathbf{Y}_t - \mathbf{1}\boldsymbol{\beta}_{0t}' - \mathbf{X}\mathbf{m}_t)$
end loop

where $\mathbf{Y}_t$ and $\mathbf{y}_{0t}$ are respectively the observations from the first and second stages of calibration. Our knowledge about the evolving relationship established in the calibration experiment at each time point is given by the posterior density $\pi(\boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2|\mathbf{Y}_t)$ which is the middle term in Equation (4.21) and found by Algorithm 5 for given values of σ_E^2 and σ_W^2. Following Algorithm 5 and using multivariate normal theory (West and Harrison 1997) we have

$$\pi(\boldsymbol{\theta}_t|\mathbf{Y}_t, \sigma_E^2, \sigma_W^2) \sim N_d(\mathbf{m}_t, \mathbf{C}_t),$$

where $\mathbf{m}_t$ is the posterior mean and $\mathbf{C}_t$ is the variance-covariance matrix of $\boldsymbol{\theta}_t$ at time t.

The first term in Equation (4.21), $f(\mathbf{y}_{0t}|x_{0t}, \boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2)$, is the likelihood function for the second stage of the calibration experiment, which provides information from the data, and $\pi(x_{0t})$ is the prior density for the unknown calibration reference x_{0t}. We wish to obtain the conditional posterior density $\pi(x_{0t}|\mathbf{y}_{0t}, \mathbf{Y}_t)$ at each time t. In order to achieve this we will have to integrate over $(\boldsymbol{\theta}, \sigma_E^2, \sigma_W^2)$.

Algorithm 6: (*SIR*) Sampling Importance Resampling

1. Draw **sample** candidates $(\boldsymbol{\Theta}^{(1)}), \ldots, (\boldsymbol{\Theta}^{(m)})$ *i.i.d.* from $g(\boldsymbol{\Theta})$
2. Calculate the standardized **importance** weights, $w(\boldsymbol{\Theta}^{(i)}) = \frac{f(\boldsymbol{\Theta}^{(i)})/g(\boldsymbol{\Theta}^{(i)})}{\sum f(\boldsymbol{\Theta}^{(i)})/g(\boldsymbol{\Theta}^{(i)})}$ for $i = 1, \ldots, m$
3. **Resample** $\boldsymbol{\Gamma}^{(1)}, \ldots, \boldsymbol{\Gamma}^{(n)}$ from $\boldsymbol{\Theta}^{(1)}, \ldots, \boldsymbol{\Theta}^{(m)}$ with replacement of probability $w(\boldsymbol{\Theta}^{(1)}), \ldots, w(\boldsymbol{\Theta}^{(m)})$ respectively.

By reducing the parameter space such that

$$\sum_{i=1}^{r} x_i = 0 \text{ and } \frac{1}{n}\sum_{i=1}^{r} x_i^2 = 1$$

which means that the columns of $\mathbf{X}$ are standardized and have average sum of squares equal one, an observation vector $\mathbf{y}_{0t}$ at time t allows an estimate of an unknown reference measurement to calculate

$$\hat{x}_{0t} = \mathbf{K}_t'\hat{\mathbf{B}}_t\hat{\boldsymbol{\Gamma}}_t^{-1}\left(\mathbf{y}_{0t} - \hat{\boldsymbol{\beta}}_{0t}\right) \tag{4.22}$$

where

$$\hat{\mathbf{K}}_t = \hat{\mathbf{B}}_t\hat{\boldsymbol{\Gamma}}_t^{-1}\hat{\mathbf{B}}_t'. \tag{4.23}$$

Given the reduced parameter space, the components of Equations (4.22) and (4.23) are derived by Algorithm 5. The estimate $\hat{\mathbf{B}}_t$ of $\mathbf{B}_t$ is the mean vector $\mathbf{m}_t$ which just contains slope elements. The estimate of the covariance matrix $\hat{\boldsymbol{\Gamma}}_t$ is given by the residual sum of products matrix $\mathbf{S}_t$ and $\hat{\boldsymbol{\beta}}_{0t} = \overline{\mathbf{Y}}_t$ at time t.

Using Equations (4.1), (4.4), (4.15), and (4.18), the likelihood function at time t is expressed as

$$f(\mathbf{y}_{0t}|x_{0t}, \boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2) \propto |\boldsymbol{\Gamma}_t|^{-1}\exp\left\{ -\tfrac{1}{2tr(\boldsymbol{\Gamma}_t)}\left[\mathbf{S}_t + r(\boldsymbol{\beta}_{0t} - \hat{\boldsymbol{\beta}}_{0t})(\boldsymbol{\beta}_{0t} - \hat{\boldsymbol{\beta}}_{0t})' + (\mathbf{B}_t - \hat{\mathbf{B}}_t)'\mathbf{X}'\mathbf{X}(\mathbf{B}_t - \hat{\mathbf{B}}_t)\right]\right\}$$

where $\hat{\boldsymbol{\beta}}_{0t}$ and $\hat{\mathbf{B}}_t$ are the dynamic parameter estimates and the residual sum product, $\mathbf{S}_t$, is used to estimate $\boldsymbol{\Gamma}_t$. Brown (1982) states that this likelihood is of the form

$$g(\boldsymbol{\beta}_{0t}, \mathbf{B}_t|\boldsymbol{\Gamma}_t)\pi(\boldsymbol{\Gamma}_t)$$

and the particular functional form suggests a natural conjugate prior distibution, such as the Normal-Inverse-Wishart distribution. Furthermore, since the columns of $\mathbf{X}$ are standardized, the prior density at time t for the unknown calibration reference is

$$\pi(x_{0t}) \propto \exp\left\{ -\frac{x_{0t}^2}{2}\right\}.$$

By integrating Equation (4.21) with respect to $\boldsymbol{\theta}_t$ for a given (σ_E^2, σ_W^2) and completing the square, the posterior density for the unknown reference measurement at time t is

$$\pi(x_{0t}|y_{0t}, \mathbf{Y}_t, \boldsymbol{\Gamma}) \sim N(\mu_{x_{0t}}, \sigma^2_{x_{0t}}), \tag{4.24}$$

with

$$\begin{aligned} \mu_{x_{0t}} &= \frac{x_{0t}}{1+\sigma^2_{Y_t}}, \\ \sigma^2_{x_{0t}} &= \frac{1}{1+\sigma^2_{Y_t}}, \end{aligned}$$

and

$$\sigma^2_{Y_t} = \text{tr}(\mathbf{Q}_t)$$

where $\sigma^2_{Y_t}$ is calculated from the trace of the one-step forecast variance-covariance matrix in Algorithm 5.

Algorithm 7: Dynamic Multi-Univariate Calibration

1. Draw M *i.i.d.* **sample** candidates for (σ_E^2, σ_W^2) from $g(\boldsymbol{\Theta})$.

2. Calibration data are fit using the Multi-univariate DLRM framework for each of the M proposal samples $(\sigma_E^{2(m)}, \sigma_W^{2(m)})$, with the prior moments for $(\boldsymbol{\theta}_0|D_0)$ as $\mathbf{m_0} = \mathbf{1_d}$ and $\mathbf{C_0} = \mathbf{100I_{(d\times d)}}$, where $\mathbf{1_d}$ is a $d-$dimensional vector of ones:

   a. Data are scaled and centered such that $\sum_{i=1}^{r} x_i = 0$ and $\frac{1}{n}\sum_{i=1}^{r} x_i^2 = 1$;

   b. Estimate $\boldsymbol{\theta}_t^{(m)}|\sigma_E^{2(m)}, \sigma_W^{2(m)}$ for the m^{th} proposal sample $\boldsymbol{\Theta}^{(m)}$ for all time t;

   c. Sample from $\pi(x_{0t}|\mathbf{y}_{0t}, \mathbf{Y}_t, \sigma_E^2, \sigma_W^2)$ given the m^{th} proposal sample $\boldsymbol{\Theta}^{(m)}$ at time t;

   d. Calculate log-likelihoods, $log[f(\boldsymbol{\Theta}^{(m)})]$, for each $(\sigma_{E_t}^{2(m)}, \sigma_{W_t}^{2(m)})$ pair.

3. Sampling Importance Resampling (SIR) is used to simulate samples of $x_{0t}|\boldsymbol{\theta}_t, \sigma_E^2, \sigma_W^2$ by accepting a subset of N from the proposal density to be distributed according to the posterior density $\pi(\sigma_E^2, \sigma_W^2|\mathbf{Y}_t)$ with candidate density $g(\boldsymbol{\Theta})$.

   a. Calculate the standardized **importance** weights, $w(\boldsymbol{\Theta}^{(1)}), \ldots, w(\boldsymbol{\Theta}^{(M)})$, where $w(\boldsymbol{\Theta}^{(m)}) = log[f(\boldsymbol{\Theta}^{(m)})] - log[g(\boldsymbol{\Theta}^{(m)})]$ for the m^{th} proposal sample;

   b. **Resample** N calibrated time series from the M proposal values with replacement given probabilities $p(\boldsymbol{\Theta}^{(m)})$ where

$$p(\boldsymbol{\Theta}^{(m)}) = \frac{e^{w(\boldsymbol{\Theta}^{(m)})}}{\sum_{j=1}^{M} e^{w(\boldsymbol{\Theta}^{(j)})}}.$$

4. Rescale calibrated time series to original scale and record summary statistics (i.e. medians and credible sets) across each time t .

To obtain posterior samples of Equation (4.24) for given σ_E^2 and σ_W^2 values, the Sampling Importance Resampling (SIR) algorithm (Rubin 1987; Smith and Gelfand 1992; Givens and Hoeting 2005; Albert 2007) is used to draw random samples from the target distribution $\pi(x_{0t}|y_{0t}, \mathbf{Y}_t, \sigma_E^2, \sigma_W^2)$ by using a joint candidate distribution $g(\boldsymbol{\Theta})$ for (σ_E^2, σ_W^2), thus deriving the posterior densities of interest. Algorithm 6 is the Sampling Importance Resampling algorithm.

Just as in the univariate case it is believed *a priori* that the observational variance σ_E^2 is greater in magnitude than the system variance σ_W^2. To enforce this belief about the variance

relationship we utilize the following prior distributions:

$$\sigma_E^2 \sim Uniform(0, \alpha_E) \tag{4.25}$$
$$\sigma_W^2|\sigma_E^2 \sim Uniform(0, \sigma_E^2), \tag{4.26}$$

where prior distributions (4.25) and (4.26) ensure the system variance to be less than the observation variance. Since these are proper prior distributions the resulting posterior distribution will also be proper. We combine prior distributions (4.25), (4.26), Algorithm 5 (MDLRM), Equation (4.24), and Algorithm 6 (SIR) together and propose the Dynamic Calibration Method in Algorithm 7.

The dynamic multi-univariate calibration approach described will be illustrated through examples followed by a comparative simulation study.

4.4 An Example

In this section, the usefulness of the dynamic multi-univariate model for calibration is illustrated through an example. The example is inspired by Landes et al. (2006), and Landes' (2010) hierarchical Bayesian statistical analysis for a calibration experiment. An experiment was conducted to enable calibration of a set of mass-produced resistance temperature devices (RTDs). The devices were placed in thermally controlled liquid baths with a precise NIST-approved thermometer where the resistance and temperature were recorded every 30 seconds. It is assumed the thermometer is accurate and each RTD responds linearly to temperature change. The goal of the Landes et al. (2006) analysis was to predict the regression parameters for an untested RTD of the same type as those used to establish the linear relationship between resistance and temperature, and derive interval estimates of temperature based on independent second-stage resistance readings.

Landes et al. (2006) considered a dozen RTDs with minimum and maximum reference temperatures of 41°F and 104°F, respectively. Conditioned on the intercept β_{0h} and slope β_{1h} given RTD h, the resistance readings from RTD h can be modeled as

$$\mathbf{Y}_{hi} = \mathbf{X}\mathbf{B}_h, + \boldsymbol{\epsilon}_{hi}, \qquad i = 1, 2, \ldots, n$$

where $\mathbf{Y}_h$ is a random vector of resistance measurements, $\mathbf{X}$ is a design matrix of fixed reference temperatures augmented with a column of 1's, $\mathbf{B}_h = [\beta_{0h}\ \beta_{1h}]'$, and $\boldsymbol{\epsilon}_i$ is an error vector with mean vector $\mathbf{0}$ and covariance matrix $\boldsymbol{\Gamma}$. Specifically, suppose that

$$\mathbf{B}_h \sim N_2[\boldsymbol{\beta}_h, \boldsymbol{\Gamma}]$$

independent of $\boldsymbol{\epsilon}_{hi}$.

For simplicity, of the 12 RTDs considered by Landes et al. (2006) only 3 will be considered here: RTD #3: RTD #4; and RTD #5. The point estimates for the $\mathbf{B}_h$'s are as follow:

$$\hat{\mathbf{B}}_3 = \begin{bmatrix} 1189 \\ 9.153 \end{bmatrix}, \qquad \hat{\mathbf{B}}_4 = \begin{bmatrix} 1166 \\ 9.441 \end{bmatrix}, \qquad \hat{\mathbf{B}}_5 = \begin{bmatrix} 1199 \\ 9.012 \end{bmatrix}.$$

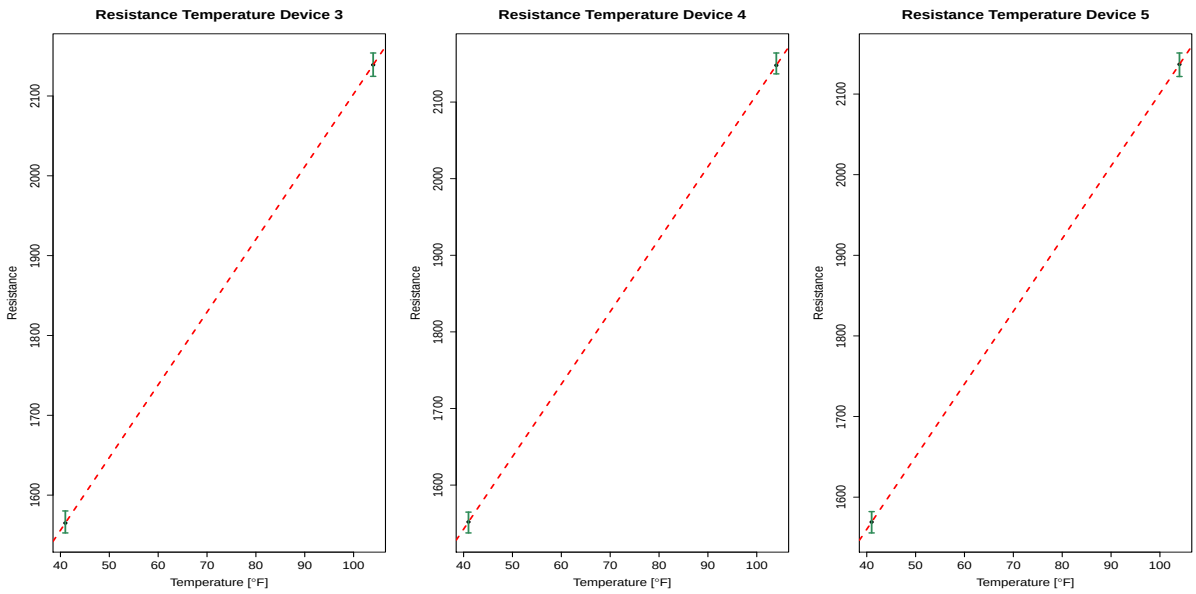


Figure 4.1: *Multi-univariate calibration example with three devices.*

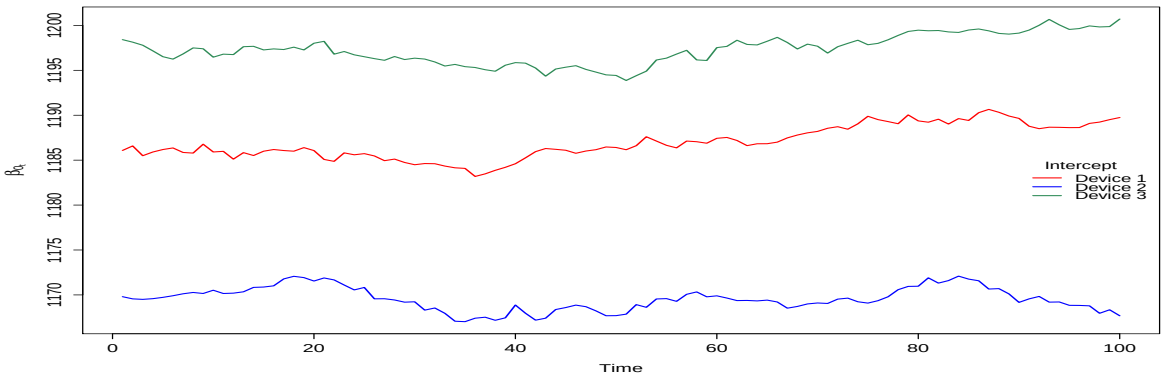
We extend the Landes et al. (2006) example by generating 100 simulated resistance measurements vectors for the 3 RTDs being considered (see Figure 4.1). The time series of resistance measurements were generated by a 50 point simple moving average (SMA_{50}) generated from 149 random draws from multivariate normal distribution with mean vector $\mathbf{B}_h$ and variance-covariance matrix

$$\mathbf{\Gamma} = \sigma^2 \begin{bmatrix} 3.14865205 & -0.0365331318 \\ -0.0365331318 & 0.0005039053 \end{bmatrix},$$

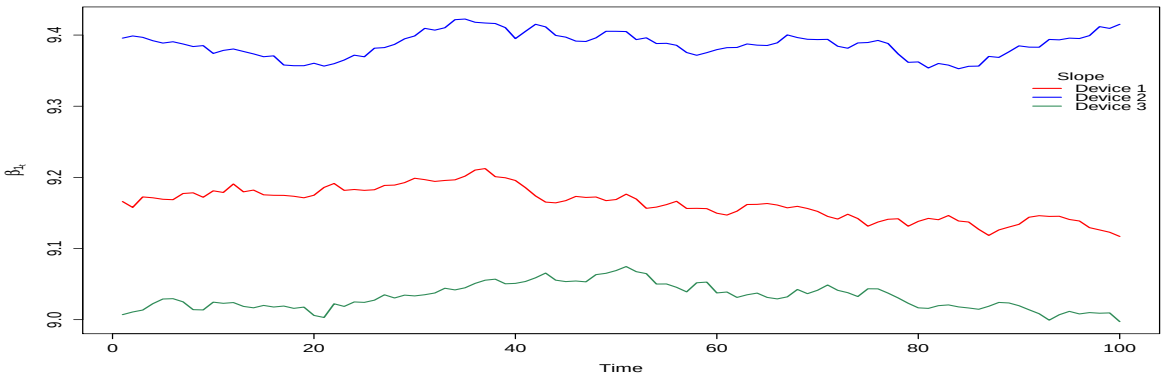
where Landes (2010) estimated $\sigma^2 = 1.287$. See Figure 4.2.

By way of example, when RTD #3 reads a resistance of $y_{0t} = 2100$ over time, the time series of the estimated temperature and 95% credible interval with respect to the drifting parameters $\mathbf{B}_{3t}$ can be found from a univariate dynamic perspective as in Figure 4.3. When calibrating an untested RTD using the dynamic multi-univariate approach, when the RTD reads a value of approximately $y_{0t} = 2100$ across each of the three tested RTD devices over time, the estimated temperature x_{0t} is given in Figure 4.5. Given that this is an example and we know the truth, $x_{0t} = 99$, performance criteria can be used as a means of evaluating the dynamic method. By incorporating information from the 3 tested RTDs, the MSE for x_{0t} calculated across the entire time series is 0.7284 with a coverage probability of 0.98 based on 95% credible intervals and an average interval width = 2.87.

Finally we consider a case when calibrating an untested RTD using the dynamic multi-univariate approach where the RTDs reads a value of approximately $y_{0t} = 1830$ across each of the three tested RTD devices over time. The target temperature x_{0t} is expected to be near the mean (= 70°F) of the two reference measurements. We see in Figure 4.5 that the



(a) *β_{0t} for the three RTDs*



(b) *β_{1t} for the three RTDs*

Figure 4.2: *Time series of β_{0t} and β_{1t} for the three RTDs*

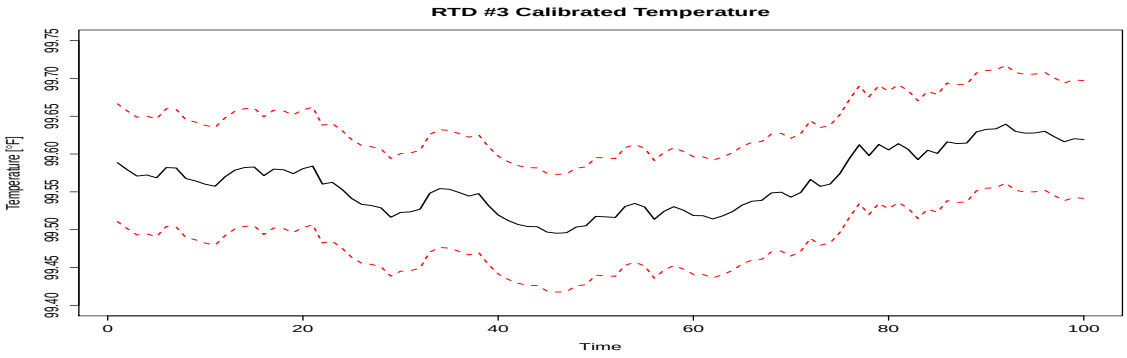


Figure 4.3: *Time series of dynamic univariate estimates of x_{0t} (Temperature [°F])*

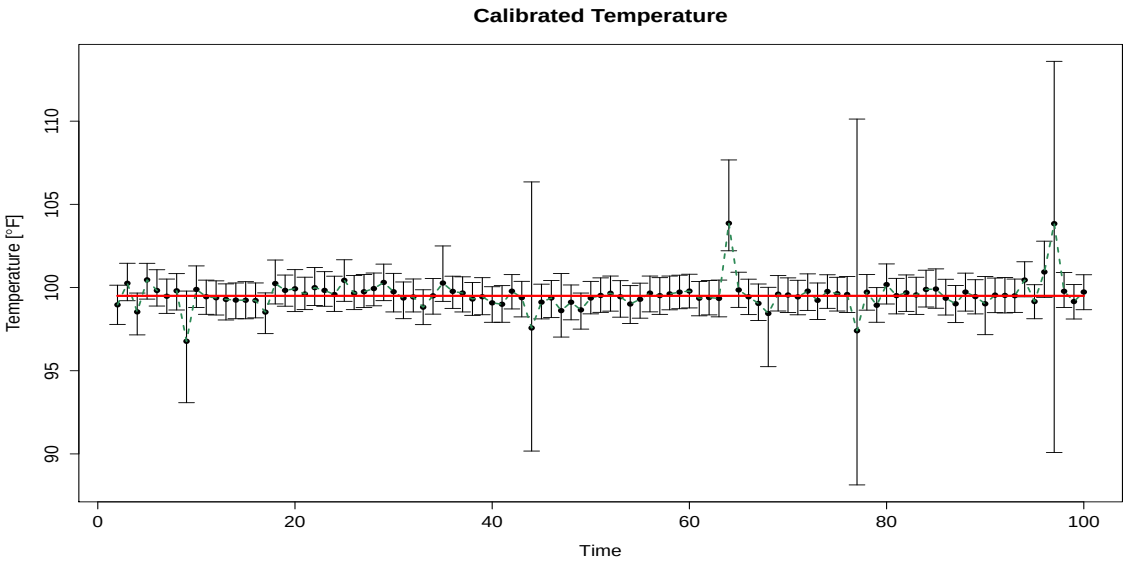


Figure 4.4: *Time series of marginal dynamic estimates of* $x_{0t}|y_{0t} = 2100$.

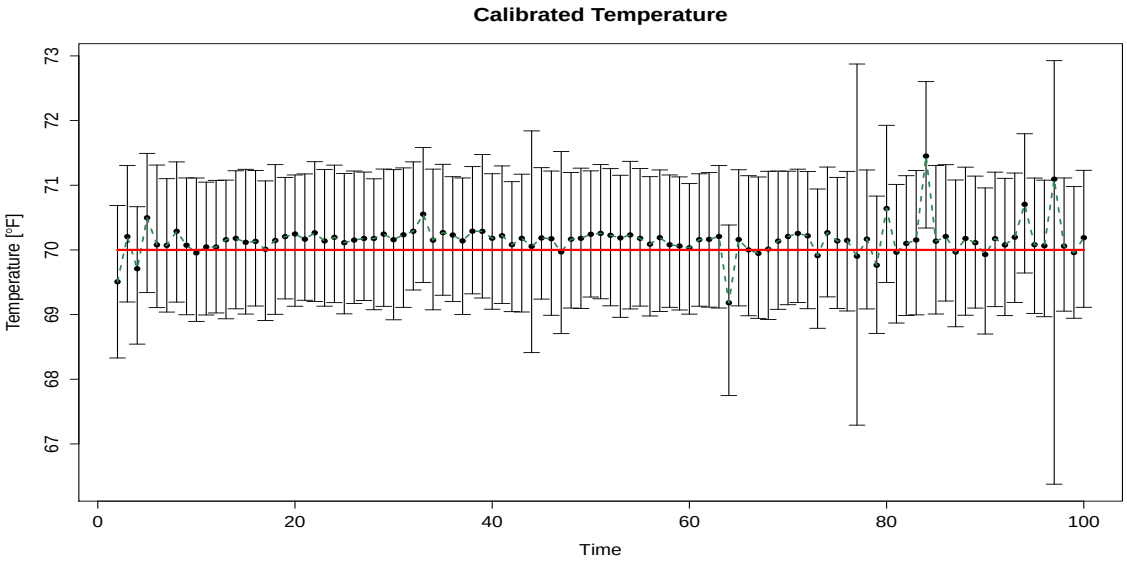


Figure 4.5: *Time series of marginal dynamic estimates of* $x_{0t}|y_{0t} = 1830$.

precision of the dynamic approach has improved compared to when the calibration target is near the boundary reference measurements. The calculated MSE for x_{0t} calculated across the entire time series is 0.08223 with a coverage probability of 0.99 based on 95% credible intervals and an average interval width = 2.23.

4.5 Simulation Study

In this section, the results of a simulation study that was conducted to evaluate the robustness of the dynamic multi-univariate calibration approach is reported. To understand what influences the calibration method, we examined three factors under three scenarios: (1) the observation variance σ_E^2; (2) the system variance σ_W^2; and (3) the time dependent marginal posterior distribution of x_{0t} corresponding to low, medium, and high values in the observed range of $\mathbf{Y}_t$. The three scenarios are as follow: (A) all devices are stable and do not exhibit systematic shocks or disturbances; (B) one of the devices experiences a systematic shock; and (C) one of the devices experiences a sinusoidal disturbances.

4.5.1 Simulation Study Design

The simulation study considers three possible system scenarios. The first scenario (A) is thought to be a perfect system that is free of random fluctuations or shocks. In other words, the variability in observed time series is due to random noise. Our second simulation (B) is designed to give insight into the behavior of the dynamic multi-univariate approach when there is a random stepped shock experienced by one of the calibration devices (see Figure 4.6). This random shock is imposed on the slope parameter β_{1t}. The third simulation design (C) is more drastic than the second design. A sinusoidal disturbance is imposed on the slope parameter β_{1t} on one of the calibration devices. This design creates a scenario where the model is disturbed for some longer period of time than just a short random shock (see Figure 4.6). For the second (B) and third (C) simulation scenarios the value γ_t is multiplied by β_{1t}, therefore modeling the affected calibration device at time t as

$$y_{rt} = \beta_{0t} + (\beta_{1t}\gamma_t)x_{rt} + \epsilon_t,$$

for the r^{th} calibration references.

The values for σ_E^2 were $(0.00001, 0.001)$ and the values for σ_W^2 were $(0.00005, 0.001)$. For each variance pair (σ_E^2, σ_W^2), the number of simulated realizations is $N = 100$. In each realization, the number of simulated time periods is $T = 1000$. The posterior densities samples drawn from $\pi(x_{0t}|y_{0t}, \mathbf{Y}_t, \mathbf{\Gamma})$ by the Dynamic Multi-Univariate Calibration Approach (Algorithm 7) is assessed via frequentist properties of the Bayesian model; the square root of the averaged mean squared errors ($RAMSE$), the average interval width ($AvIW$), and

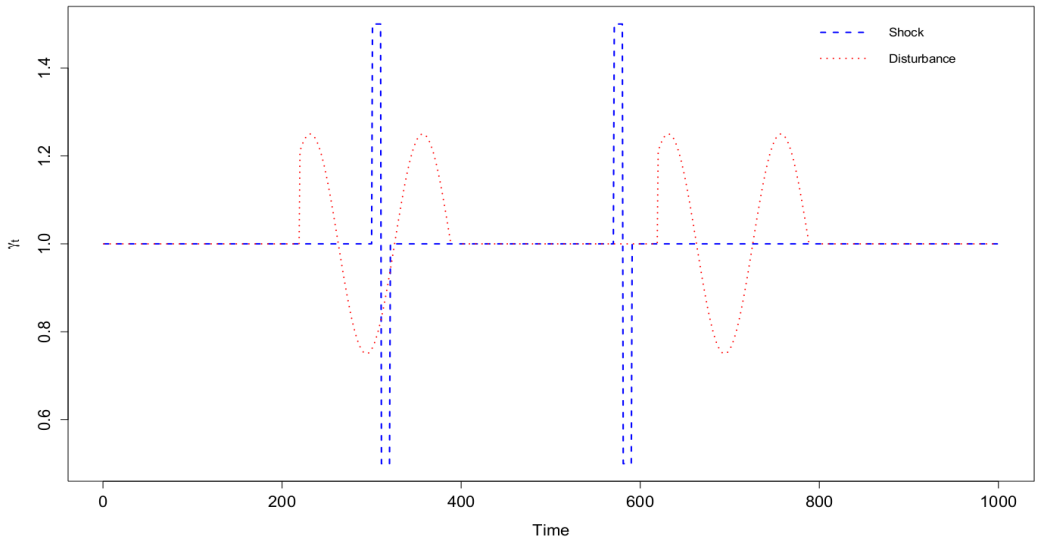


Figure 4.6: γ_t*: Short stepped shock and long sinusoidal disturbance.*

the average coverage probability ($AvCP$):

$$RAMSE = \left[\frac{1}{N}\sum_{j=1}^{N} MSE_j\right]^{\frac{1}{2}},$$

where

$$MSE_j = \frac{1}{T}\sum_{t=1}^{T}(\hat{x}_{0t} - x_{0t})^2;$$

and

$$AvIW = \frac{1}{N}\sum_{j=1}^{N} IW_j.$$

where

$$IW_j = \frac{1}{T}\sum_{t=1}^{T}(x_{0t}^U - x_{0t}^L).$$

Note that if x_{0t}^L is the 0.025 posterior quantile for x_{0t}, and x_{0t}^U is the 0.975 posterior quantile for x_{0t}, where x_{0t} is the true value of the calibration target from the second stage of experimentation, then (x_{0t}^L, x_{0t}^U) is a 95% credible interval.

Using the credible interval above we defined the coverage probability (CP) which is calculated as such

$$CP_j = \frac{1}{T}\sum_{t=1}^{T}\psi_t$$

where

$$\psi_t = P[x_{0t}^L < x_{0t} < x_{0t}^U] = \begin{cases} 0 & \text{if } x_{0t} \notin (x_{0t}^L, x_{0t}^U); \\ \\ 1 & \text{if } x_{0t} \in (x_{0t}^L, x_{0t}^U). \end{cases}$$

The average coverage probability ($AvCP$) is calculated by averaging across the number of replications in the simulation study, where

$$AvCP = \frac{1}{N} \sum_{j=1}^{N} CP_j.$$

The data are generated according to the following model described in Section 4.3. The design matrix $\mathbf{X}$ consist of reference values [70, 140] augmented with a column of 1's. The vector of regression parameters, $\mathbf{B}_{ht}$, for each hypothetical device h are random draws from $h = 1, 2, 3$ multivariate normal distribution with mean vector

$$\mathbf{B}_{ht} = \begin{bmatrix} 1200 \\ 10 \end{bmatrix}$$

and variance-covariance matrix, $\mathbf{W} = \sigma_W^2 \left[\mathbf{X}'\mathbf{X}\right]^{-1}$ for $t = 1, \ldots, T$. For each of the h devices at time t, the random multivariate error vector is

$$\boldsymbol{\epsilon}_{ht} \sim N_r[\mathbf{0}, \sigma_E^2 \mathbf{I}]. \tag{4.27}$$

The marginal dynamic estimates posterior distribution of x_t are based on three values of y_{0t}. The realizations have mean equal: $y_{0t} = 1550$; $y_{0t} = 1850$; and $y_{0t} = 2150$ across each of the calibration devices with innovation error variance σ_E^2. The dynamic approach denoted in the tables (DC) will be compared the frequentist inverse multivariate approach (FC) and the Bayesian multivariate approach (BC) of Brown (1982). Due to singularity concerns, the g-inverse $\hat{\Gamma}^-$ will be inserted in place of $\hat{\Gamma}^-$ for (FC) and (BC). The data sets were generated in R version 3.1.2 (Pumpkin Helmet) (R Core Team 2013).

4.5.2 Simulation Study Results

The results of the simulation study are summarized in the following tables. The following analysis is based on 3600 data sets on the dynamic multi-univariate approach described in Section 4.3 and two static multivariate calibration methods from Section 4.2. The first static multivariate calibration method is Brown (1982) inverse frequentist approach described by Equation 4.7. Brown's (1982) multivariate version of Hoadley (1970) Bayesian approach (*Theorem* 1) is the second static multivariate calibration method compared through the simulation study.

The aforementioned calibration methods are assessed for three cases (i.e. Constant, Stepped, Sinusoidal) across four combination of the variance pair (σ_E^2, σ_W^2) for three levels of

the calibration observation $y_{0t} = 1550$; $y_{0t} = 1850$; and $y_{0t} = 2150$. The results for the case where each of the devices in the multi-univariate calibration experiment is free of systematic disturbances is given in Table 4.1. The case when one of the calibration devices experience a systematic stepped noise is provided in Table 4.2. The results for the sinusoidal disturbance case is given in Table 4.3.

For the case without systematic influence on one of the devices (in Table 4.1), the

Table 4.1: *Main results of the simulation study for the case without a system fluctuation: RAMSE, AIW and AvCP are given for the dynamic multi-univariate calibration estimator (DC), frequentist estimator (FC), and the Bayesian estimator (BC) under* (σ_E^2, σ_W^2)*. Results are provided for* $y_{0t} = 1550$*,* $y_{0t} = 1850$*, and* $y_{0t} = 2150$*.*

					$\mathbf{y_{0t} = 1550}$					
		RAMSE			**AIW**			**ACP**		
σ_W^2	σ_E^2	DC	FC	BC	DC	FC	BC	DC	FC	BC
0.00005	0.00001	0.501	0.501	44.590	2.090	100.852	25.412	1.000	1.000	0.954
	0.001	0.502	0.510	459.259	2.095	100.852	25.412	1.000	1.000	0.951
0.001	0.00001	0.501	0.502	147.798	2.095	100.852	25.412	1.000	1.000	0.951
	0.001	0.502	0.512	411.454	2.098	100.852	25.412	1.000	1.000	0.951
					$\mathbf{y_{0t} = 1850}$					
0.00005	0.00001	0.041	0.038	511.493	2.094	100.852	25.412	1.000	1.000	0.947
	0.001	0.041	0.079	10552.901	2.094	100.852	25.412	1.000	1.000	0.951
0.001	0.00001	0.041	0.039	88.910	2.093	100.852	25.412	1.000	1.000	0.949
	0.001	0.041	0.080	70.009	2.094	100.852	25.412	1.000	1.000	0.948
					$\mathbf{y_{0t} = 2150}$					
0.00005	0.00001	0.081	0.080	375.365	2.098	100.852	25.412	1.000	1.000	0.949
	0.001	0.081	0.128	201.704	2.092	100.852	25.412	1.000	1.000	0.953
0.001	0.00001	0.081	0.082	541.780	2.096	100.852	25.412	1.000	1.000	0.949
	0.001	0.081	0.129	222.785	2.095	100.852	25.412	1.000	1.000	0.949

dynamic calibration approach and the frequentist approach are near comparable in terms of RAMSE for $\mathbf{y_{0t} = 1550}$ and $\mathbf{y_{0t} = 1550}$. When $\mathbf{y_{0t} = 1550}$, the targeted reference measurement falls outside of the range of initial reference measurements of $41°K$ and $104°K$. Brown (1982) Bayesian multivariate approach is clearly not appropriate for this sort of analysis. It performs poorly compared to the dynamic approach as well as the frequentist approach. The dynamic approach consistently has an interval width smaller than this for the frequentist model (FC) and the Bayesian model (BC). This average interval width gives insight into the precision of the method. It seems as though when the devices are free of disturbances over the given time period, the dynamic approach and the frequentist approach are very similar.

Table 4.2: *Main results of the simulation study for the case with a system stepped fluctuation: RAMSE, AIW and AvCP are given for the dynamic multi-univariate calibration estimator (DC), frequentist estimator (FC), and the Bayesian estimator (BC) under* (σ_E^2, σ_W^2)*. Results are provided for* $y_{0t} = 1550$*,* $y_{0t} = 1850$*, and* $y_{0t} = 2150$*.*

					$\mathbf{y_{0t} = 1550}$					
		RAMSE			**AIW**			**ACP**		
σ_W^2	σ_E^2	DC	FC	BC	DC	FC	BC	DC	FC	BC
0.00005	0.00001	0.501	3.740	69.339	2.094	100.852	25.412	1.000	1.000	0.936
	0.001	0.502	3.743	67.565	2.091	100.852	25.412	1.000	1.000	0.934
0.001	0.00001	0.502	3.740	79.565	2.091	100.852	25.412	1.000	1.000	0.928
	0.001	0.502	3.743	71.987	2.095	100.852	25.412	1.000	1.000	0.931
					$\mathbf{y_{0t} = 1850}$					
0.00005	0.00001	0.041	6.741	61.334	2.092	100.852	25.412	1.000	1.000	0.932
	0.001	0.041	6.742	132.471	2.092	100.852	25.412	1.000	1.000	0.932
0.001	0.00001	0.041	6.741	231.452	2.095	100.852	25.412	1.000	1.000	0.932
	0.001	0.041	6.742	778.111	2.098	100.852	25.412	1.000	1.000	0.928
					$\mathbf{y_{0t} = 2150}$					
0.00005	0.00001	0.081	9.845	99.076	2.092	100.852	25.412	1.000	0.980	0.932
	0.001	0.081	9.845	542.313	2.095	100.852	25.412	1.000	0.980	0.924
0.001	0.00001	0.081	9.844	777.577	2.096	100.852	25.412	1.000	0.980	0.931
	0.001	0.081	9.847	103.899	2.093	100.852	25.412	1.000	0.980	0.928

From a review of the results in Tables 4.2 and 4.3 we see that the dynamic approach (DC) is invariant to the stepped and sinusoidal fluctuations imposed on one of the devices. The results for the dynamic approach performance in Table 4.2 and Table 4.3 is consistent to the results given in Table 4.1. The frequentist (FC) and Bayesian (BC) approaches are unable to adapt to imposed stepped fluctuation as well as the dynamic method (DC). The dynamic method is clearly superior to the frequentist (FC) and Bayesian (BC) approached to the multi-univariate calibration problem.

Table 4.3: *Main results of the simulation study for the case with a system sinusoidal fluctuation: RAMSE, AIW and AvCP are given for the dynamic multi-univariate calibration estimator (DC), frequentist estimator (FC), and the Bayesian estimator (BC) under* (σ_E^2, σ_W^2)*. Results are provided for* $y_{0t} = 1550$*,* $y_{0t} = 1850$*, and* $y_{0t} = 2150$*.*

					$\mathbf{y_{0t} = 1550}$					
		RAMSE			**AIW**			**ACP**		
σ_W^2	σ_E^2	DC	FC	BC	DC	FC	BC	DC	FC	BC
0.00005	0.00001	0.501	8.221	208.675	2.096	100.852	25.412	1.000	1.000	0.822
	0.001	0.501	8.221	72.217	2.094	100.852	25.412	1.000	1.000	0.824
0.001	0.00001	0.501	8.221	122.166	2.092	100.852	25.412	1.000	1.000	0.831
	0.001	0.502	8.223	115.276	2.092	100.852	25.412	1.000	1.000	0.831
					$\mathbf{y_{0t} = 1850}$					
0.00005	0.00001	0.041	15.174	401.370	2.089	100.852	25.412	1.000	1.000	0.721
	0.001	0.041	15.172	46.330	2.091	100.852	25.412	1.000	1.000	0.718
0.001	0.00001	0.041	15.174	106.563	2.095	100.852	25.412	1.000	1.000	0.717
	0.001	0.041	15.178	315.720	2.091	100.852	25.412	1.000	1.000	0.719
					$\mathbf{y_{0t} = 2150}$					
0.00005	0.00001	0.081	22.135	254.994	2.099	100.852	25.412	1.000	0.919	0.679
	0.001	0.081	22.136	50.352	2.093	100.852	25.412	1.000	0.919	0.676
0.001	0.00001	0.081	22.136	131.790	2.096	100.852	25.412	1.000	0.919	0.685
	0.001	0.081	22.134	175.016	2.090	100.852	25.412	1.000	0.919	0.682

4.6 Conclusion

By comparing Tables 4.1, 4.2 and 4.3 we see the dynamic approach to the multi-univariate calibration problem is very good option when the measurements are correlated values taken over time. The given scenarios illustrates different situations that may occur in practice when calibrating sensitive devices. It is clear that static methods may not be appropriate for calibration experiments designed in the manner depicted in this work, therefore an approach that is adaptive to various calibration schemes and robust under systematic fluctuations. We demonstrated that the proposed approach is invariant to disturbances that may be happening to a single calibration device while estimating the value of the desired calibration target. Furthermore, we can see that our method allowed for us to detect the evolving relationships between the calibration references and the measurable output which were difficult to capture using the other methods. We are not surprised by the results due to the fact that the dynamic method is designed purposefully learn the behavior of the data thus use *a priori* information at each time step to predict the most likely model parameters therefore increasing the precision of the estimator.

4.7 References

[1] Broekhuizen, S., and Maaskamp, F. (1979). Age determination in the European hare (Lepus europaeus Pallas) in the Netherlands. *Zeitschrift frü Säugetierkunde.* **44**(3), 162-175.

[2] Brown, G. H. (1979). An optimization criterion for linear inverse estimation. *Technometrics*, **21**(4), 575-579.

[3] Brown, P. J. (1982). Multivariate calibration. *Journal of the Royal Statistical Society B.* **44**, 287-321.

[4] Brown, P. J. (1993). *Measurement, Regression, and Calibration.* Oxford University Press.

[5] Brown-1987 Brown, P. J. and Sundberg, R. (1987). Confidence and conflict in multivariate calibration. *Journal of the Royal Statistical Society B.* **49**, 46-57.

[6] Cervenka, P. O., and Massa, L. (1994). Nonlinear Calibration of an Infrared Radiometer (No. CARDIVNSWC-TR-94/024). NAVAL SURFACE WARFARE CENTER CARDEROCK DIV BETHESDA MD.

[7] duPlessis-1995 du Plessis, J. L., and van der Merwe, A. J. (1995). A Bayesian Approach to Multivariate and Conditional Calibration. *Computational Statistics & Data Analysis.* **19**(5), 539-552.

[8] Eisenhart, C. (1939). The interpretation of certain regression methods and their use in biological and industrial research. *Annals of Mathematical Statistics*. **10**, 162-186.

[9] Fearn-1983 Fearn, T. (1983). Misuses of Ridge Regression in the Calibration of a Near Infra-red Reflectance Instrument. *Appl. Statist.* **32**, 73-79.

[10] Fox-1989 Fox, D. R. (1989). Conditional Multivariate Calibration. *Communications in Statistics-Theory and Methods*. **18**(6), 2311-2330.

[11] Gujral, P., Amrhein, M., and Bonvin, D. (2009). Drift correction in multivariate calibration models using on-line reference measurements. *Analytica chimica acta*. **642**(1), 27-36.

[12] Hoadley, B. (1970). A Bayesian look at Inverse Linear Regression. *Journal of the American Statistical Association*. **65**, 356-369.

[13] Hunter, W., and Lamboy, W. (1981). A Bayesian Analysis of the Linear Calibration Problem. *Technometrics*. **23**, 323-328.

[14] Kalbfleisch-1970 Kalbfleisch, J. D. and Sprott, D. A. (1970). Applications of Likelihood Methods to Models Involving Large Numbers of Parameters (with discussion). *J. R. Statist. Soc. B* **32**, 175-208.

[15] Landes, R. D. (2010). Hierarchical Bayesian Calibration of Untested Devices. *Communications in Statistics-Simulation and Computation*. **39**(7), 1351-1364.

[16] Landes, R. D., Loutzenhiser, P. G., and Vardeman, S. B. (2006). Hierarchical Bayesian statistical analysis for a calibration experiment. *Instrumentation and Measurement, IEEE Transactions on*, **55**(6), 2165-2171.

[17] Naes-1986 Naes, T., Irgens, C. and Martens, H. (1986). Comparison of Linear Statistical Methods for Calibration of NIR Instruments. *Appl. Statist.* **35**,195-206.

[18] Nishii, R., and Krishnaiah, P. R. (1988). On the moments of classical estimates of explanatory variables under a multivariate calibration model. *Sankhyā: The Indian Journal of Statistics, Series A*, 137-148.

[19] Oman, S. D., and Srivastava, M. S. (1996). Exact mean squared error comparisons of the inverse and classical estimators in multi-univariate linear calibration. *Scandinavian Journal of Statistics*, 473-488.

[20] Oman, S. D., and Wax, Y. (1984). Estimating fetal age by ultrasound measurements: an example of multivariate calibration. *Biometrics*, 947-960.

[21] Osborne, C. (1991). Statistical Calibration: A Review. *International Statistical Review*. **59**(3), 309-336.

[22] Özyurt, Ö., and Erar, A. (2003). Multivariate Calibration in Linear Regression and its Application. *Hacettepe Journal of Mathematics and Statistics*. **32**, 53-63.

[23] Petris, G., Petrone, S., and Campagnoli, P. (2009). Dynamic linear models with R. Springer.

[24] R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

[25] Sundberg, R., and Brown, P. J. (1989). Multivariate calibration with more variables than observations. *Technometrics*. **31**(3), 365-371.

[26] Srivastava, M. S. (1995). Comparison of the inverse and classical estimators in multi-univariate linear calibration. *Communications in Statistics-Theory and Methods*. **24**(11), 2753-2767.

[27] Takeuchi, H. (1997). A generalized inverse regression estimator in multi-univariate linear calibration. *Communications in Statistics-Theory and Methods*. **26**(11), 2645-2669.

Chapter 5

Closing Remarks and Future Work

Several novel approaches to the statistical calibration problem have been presented in this dissertation. We showed through simulation that the dynamic approach has its advantages over the traditional static methods. Clearly, if the relationship between the predictor variable and the response variable in the first stage of calibration experiment is known to be stable then the traditional methods should be used. The dynamic approach is computationally expensive if the calibration environment is not likely to change statistically over time. If it is required to estimate repeated calibrated values over time then the dynamic approach should be considered. The methods proposed in this dissertation have been showed to promising in cases when the variance ratio is high.

The nonlinear case examined in Chapter 3 provided great insight in to the flexibility of the dynamic calibration approach. This model was found to be a robust under various scenarios when approximating the posterior distribution of an unknown reference measurement when there is nonlinearity in the covariate. The dynamic nonlinear approach was adaptable given the number of reference measurements and the placement of these references as they define the curvature. Our approach successfully combines Bayesian time series analysis, nonlinear calibration, and sampling theory.

The multi-univariate approach was a clear extension to the calibration problem. We demonstrated that the dynamic multi-univariate calibration approach provided consistent posterior estimates even when one of the devices fall victim to instabilities in the system such as systematic gain fluctuations. Due to the recursive Kalman Filtering estimation process the dynamic multi-univariate approach is not restricted to the normal multivariate assumptions. It was made clear that the static methods may not be appropriate for calibration experiments designed in the manner illustrated in this dissertation.

In the methods presented in this dissertation, we have considered only linear regression models with a single covariate. Future work for the dynamic model proposed in Chapter 2 could be considered from a multiple linear regression framework. Through this extension to the linear calibration problem one could study how the method performs when the response is univariate and multivariate in the covariates. It may be interest to explore models that

can be framed in this manner. The framework of the nonlinear approach in Chapter 3 could be extended perform calibration when the model is non-linear in the parameters. François et al. (2004) looked at the calibration when using exponential and four parameter logistic (4pl) models, while Kirkup and Mulholland (2004) not only considered the quadratic model but also discussed a model that was nonlinear in the parameters. It may be of interest to further investigate the performance of the the proposed methods under these extensions.

Appendix

The section contains the R code that was used for the simulation study conducted in the dissertation.

A Dynamic Linear Calibration R Code

A.1 Data Generating Code

The data generating R script file:

```
#--------------------------
# Design Matrix of 2 Known
# Reference Temperatures
#--------------------------

n1 = 1074 #Initial Simulated Values
T1 = 20   #Reference Temp 1
T2 = 100 #Reference Temp 2

X  <- matrix(c(1,1,T1,T2), 2, 2, byrow = F) #Design Matrix

#--------------------------
# Design Matrix of 5 Known
# Reference Temperatures
#--------------------------

# n1 = 1074 #Initial Simulated Values
# T1 = 20 #Reference Temp 1
# T2 = 40 #Reference Temp 2
# T3 = 60 #Reference Temp 3
# T4 = 80 #Reference Temp 4
# T5 = 100#Reference Temp 5
# X  <- matrix(c(1,1,1,1,1,T1,T2,T3,T4,T5), 5, 2, byrow = F)

#--------------------------
# Load in theta_t
# theta = [12.7434, 0.02655]
#
# theta_1 -> sigma2_w = 0.00001
```

```
# theta_5 -> sigma2_w = 0.00005
#--------------------------

theta_1 <- read.table("/home/riversdl/Needed_Files/theta_t(1).txt")
theta_5  <- read.table("/home/riversdl/Needed_Files/theta_t(5).txt")

#--------------------------
# Load in g_t (Gain Noise)
#
# goft[,1] = constant
# goft[,2] = sinusoidal
# goft[,3] = stepped
#--------------------------

goft <- read.csv("/home/riversdl/Needed_Files/goft.csv")

#--------------------------
# Generate 2 Vectors of
# Observations given X*Theta_t,
# sigma2_E = [0.0001, 0.001, 0.01]
# and gamma = [0.025, 0.0025, 0.00025]
#--------------------------

sig1  = 0.0001  #[0.0001, 0.001, 0.01]                        #Sig_E
gam   = 0.25    #[0.25 for Sinusoidal, 0.025 for Stepped]    #Gamma (gain)
err   = rnorm(dim(theta_1)[1],sd = sqrt(sig1))               #Noise term

Yt1   <- theta_1[,1] + (t(X[1,2])
Yt2   <- theta_1[,1] + (t(X[2,2])

Yt    <- rbind(Yt1, Yt2)

#--------------------------
# Generate 5 Vectors of
# Observations given X*Theta_t,
# sigma2_E = [0.0001, 0.001, 0.01]
# and gamma = [0.025, 0.0025, 0.00025]
#--------------------------

# sig1  = 0.0001    #[0.0001, 0.001, 0.01]                     #Sig_E
# gam   = 0.25      #[0.025, 0.0025, 0.00025]                  #Gamma (gain)
# err   = rnorm(dim(theta_5)[1],sd = sqrt(sig1))               #Noise term
```

```
#
#
# Yt1   <- theta_5[,1] + (t(X[1,2])
# Yt2   <- theta_5[,1] + (t(X[2,2])
# Yt3   <- theta_5[,1] + (t(X[3,2])
# Yt4   <- theta_5[,1] + (t(X[4,2])
# Yt5   <- theta_5[,1] + (t(X[5,2])
#
# Yt    <- rbind(Yt1, Yt2, Yt3, Yt4, Yt5)

#------------------------------
# Load the true calibration
# brightness temperature x0t
#
# Load the observed voltage
# measurement y0t
#------------------------------

x0t_50 <- read.table("/home/riversdl/Needed_Files/x0t_50.txt", quote="\"")

y0t_50  <- read.table("/home/riversdl/Needed_Files/y0t_50.txt", quote="\"")
```

A.2 Dynamic Linear Calibration Code

The linear Dynamic Linear Calibration R script file:

```
dlm.Est <- function(X, Y, m.0, C.0, W.t, E.t, y0){


  tot = dim(X)[1]                           #Total number of references
  Sxx = t(X[,2] - mean(X[,2]))^2            #Reference sum of squares
  a   = sqrt(tot/Sxx)                       #Scaling coefficient
  X   = a*(X[,2] - mean(X[,2]))
  X   = matrix(X,tot,1)                     #Scaled reference measurements
  X <- cbind(1,X)

  Y <- na.omit(Y)
  n <- dim(Y)[1]

  R.t <- array(0,c(dim(m.0)[1],dim(m.0)[1],n))#R.t=C.t+W.t
  a.t <- matrix(0,dim(m.0)[1],n)              #a.t=m.(t-1)
```

```
f.t <- matrix(0,dim(X)[1],n)                 #f.t=t(X.t)*a.t
Q.t <- array(0,c(dim(X)[1],dim(X)[1],n))     #Q.t=t(X.t)*R.t*X.t+S.(t-1)
m.t <- matrix(0,dim(m.0)[1],n)               #m.t=a.t+A.t*e.t
C.t <- array(0,c(dim(m.0)[1],dim(m.0)[1],n))#C.t=(R.t*E.t)/Q.t
A.t <- array(0,c(dim(m.0)[1],dim(X)[1],n))  #A.t=(R.t1*X.t)/Q.t
e.t <- matrix(0,dim(X)[1],n)                 #e.t=Y.t-f.t

known.sigma.sq <- matrix(0,n)
post.mean      <- matrix(0,n)
post.var       <- matrix(0,n)

x01  <- matrix(0,n)
x02  <- matrix(0,n)

SSy <- matrix(0,n)
SSb <- matrix(0,n)
###################################
####  Initialize the vectors  #####
###################################

R.t[,,1] <- C.0 + W.t
a.t[,1]  <- m.0
f.t[,1]  <- X %*% a.t[,1]
Q.t[,,1] <- X %*% R.t[,,1] %*% t(X) + E.t
A.t[,,1] <- (R.t[,,1] %*% t(X)) %*% solve(Q.t[,,1])
e.t[,1]  <- t(Y[1,]) - mean(t(Y[1,])) - f.t[,1]
m.t[,1]  <- a.t[,1] + diag(c(1,1),2,2) %*% (A.t[,,1] %*% e.t[,1])
C.t[,,1] <- R.t[,,1] - A.t[,,1] %*% Q.t[,,1] %*% t(A.t[,,1])

SSb[1]   <- t(m.t[,1] - m.0) %*% (m.t[,1] - m.0)
SSy[1]   <- t(e.t[,1]) %*% e.t[,1]

################################
##  Calibration Distribution  ##
################################

   mu0     = 0
   tau.sq0 = 1
   n0      = 1
   known.sigma.sq[1] = tr(Q.t[,,1])

   post.mean[1] <- (mu0 * known.sigma.sq[1] + tau.sq0 *
```

```
      (y0[1]- mean(t(Y[1,])))/rnorm(1,m.t[2,1],sqrt(C.t[2,2,1])))/
     (n0 * tau.sq0 + known.sigma.sq[1])

     post.var[1]  <- (known.sigma.sq[1] * tau.sq0)/
      (n0 * tau.sq0 + known.sigma.sq[1])

     x02[1]    <- rnorm(1,post.mean[1], sqrt(post.var[1]))

     x01[1]    <- (y0[1]- mean(t(Y[1,])))/
      rnorm(1,m.t[2,1], 10*sqrt(C.t[2,2,1]))

  ###############################
  ####  Bayesian Updating    ####
  ###############################

  for(i in 2:n){

    R.t[,,i] <- C.t[,,i-1] + W.t
    a.t[,i]  <- m.t[,i-1]
    f.t[,i]  <- X %*% a.t[,i]
    Q.t[,,i] <- X %*%  R.t[,,i] %*% t(X) + E.t
    A.t[,,i] <- (R.t[,,i] %*% t(X)) %*% solve(Q.t[,,i])
    e.t[,i]  <- t(Y[i,]) - mean(t(Y[i,])) - f.t[,i]
    m.t[,i]  <- a.t[,i] + diag(c(1,1),2,2) %*% (A.t[,,i] %*% e.t[,i])
    C.t[,,i] <- R.t[,,i] - A.t[,,i] %*% Q.t[,,i] %*% t(A.t[,,i])

    SSb[i]   <- t(m.t[,i] - m.t[,i-1]) %*% (m.t[,i] - m.t[,i-1])
    SSy[i]   <- t(e.t[,i]) %*% e.t[,i]

    ###############################
    ## Calibration at point y0 ##
    ###############################

    known.sigma.sq[i] = tr(Q.t[,,i])

    post.mean[i] <- (mu0 * known.sigma.sq[i] + tau.sq0 *
     (y0[i]- mean(t(Y[1:i,])))/
     rnorm(1,m.t[2,i],sqrt(C.t[2,2,i])))
/(n0 * tau.sq0 + known.sigma.sq[i])

    post.var[i]  <- (known.sigma.sq[i] * tau.sq0)/
     (n0 * tau.sq0 + known.sigma.sq[i])
```

```
    x02[i]    <- rnorm(1,post.mean[i], sqrt(post.var[i]))

    x01[i]    <- (y0[i] - mean(Y[1:i,]))/
     rnorm(1, m.t[2,i], 10*sqrt(C.t[2,2,i]))
  }

  ####################
  ## Sum of Squares ##
  ####################

  tb1 <-  sum(SSb[1:n])/n  #11
  syy <-  sum(SSy[1:n])/n  #11

  param <- list(x01=x01, x02=x02, tb1=tb1, SSy=syy, a=a)
  return(param)

}
```

A.3 Sampling Importance Resampling Wrapper Code

The Sampling Importance Resampling Wrapper `R` script file:

```
#--------------------
# Load library
# of needed packages
#--------------------

library(geoR)
library(psych)
library(Hmisc)

setwd("/home/riversdl/Statistical_Calibration")

#--------------------
# Simulation Counter
#--------------------

Sim = 100
```

```
for(c in 1:Sim){

#--------------------
# Sourced Files
#--------------------

source("Data_Gen1.R")      #Data_Gen_1 (2 ref) or Data_Gen_2 (5 ref)
source("Static_Models.R")  #Statice Frequentist and Bayesian Approaches
source("Sim_DLM(Bayes).R") #Dynamic Calibration Methods

var.est <- 1

N     = 1000   # Number of time periods
alpha = 0.025  # [0.005,0.995] [0.025,0.975]  [0.05, 0.95]
#---------------------------------------
# Load the true calibration
# brightness temperature x0t
#
# Load the observed voltage
# measurement y0t
#---------------------------------------

true.x0 = x0t_50[,c]
y0t     = y0t_50[,c]

#--------------------------------
# Run Static Calibration Models
#--------------------------------

    #----------------------------------------
    # Data must be structured for static models
    #----------------------------------------

X11 <- cbind(rep(X[1,2], length(Yt[1,])), Yt[1,])
X12 <- cbind(rep(X[2,2], length(Yt[2,])), Yt[2,])
# X13 <- cbind(rep(X[3,2], length(Yt[3,])), Yt[3,])
# X14 <- cbind(rep(X[4,2], length(Yt[4,])), Yt[4,])
# X15 <- cbind(rep(X[5,2], length(Yt[5,])), Yt[5,])
# data <- rbind(X11,X12,X13,X14,X15)
data <- rbind(X11,X12)
X1 <- data[,1]
Y1 <- data[,2]
```

```
x = X1
y = Y1
df = data.frame(x,y)

#-------------------------------------------------------
# "class" is Eisenhart's Classical Calibration Approach
# "inver" is Krutchkoff's Inverse Calibration Approach
# "hoadl" is Hoadley's Bayesian Approach
# "huntl" is Hunter & Lamboy's Bayesian Approach
#-------------------------------------------------------

class <- calib1(X1, Y1, alpha, y0t)
inver <- calib2(X1, Y1, alpha, y0t)
hoadl <- calib3(df,     alpha, y0t)
huntl <- calib4(X1, Y1, alpha, y0t)

#--------------------------------
# Sampling Importance Resampling
#--------------------------------

simval = 10000
obs_prior1 <- matrix(0,simval,1)
sys_prior3 <- matrix(0,simval,1)
sig_value  <- matrix(0,simval,2)
p.weight   <- matrix(0,simval,1)
candidate  <- matrix(0,simval,1)
weights    <- matrix(0,simval,1)
exp.weights<- matrix(0,simval,1)

#-----------------------
# Use for Deterministic
# Approach to Calibration
#-----------------------

#  results    <- matrix(0,simval,3003)

#-----------------------
# Use for Bayesian
# Approach to Calibration
#-----------------------

  results    <- matrix(0,simval,2003)
```

```
#---------------------------------------
# Candidate Distribution
# for Variance Pairs (sigma2_E, sigma2_W)
#---------------------------------------

for(i in 1:simval){

  obs_prior1[i] <- runif(1,0,var.est)
  sys_prior3[i] <- runif(1,0,obs_prior1[i])

  sig_value[i,] <- c(obs_prior1[i], sys_prior3[i])

}

#-----------------------------------------
# Run the Dynamic Calibration Algorithm
#     and derive importance weights
#-----------------------------------------

for(i in 1:dim(sig_value)[1]){

  m.0 = matrix(c(10,0.05),2,1)
  C.0 = diag(rep(100,2))

  letsgo <-  dlm.Est(X, t(Yt), m.0, C.0,
   sig_value[i,2]*diag(dim(m.0)[1]),
sig_value[i,1]*diag(dim(X)[1]), y0 = y0t)

  a1    <-  log(1/var.est) - log(sig_value[i,1])  #log(1/var.est)
  like1 <- -(10/2)*log(sig_value[i,1])-(1/(2*sig_value[i,1]))*letsgo$SSy
  like2 <- -(10/2)*log(sig_value[i,2])-(1/(2*sig_value[i,2]))*letsgo$tb1

  p.weight[i] <- a1 + like1 + like2

  candidate[i] <-

    dunif(sig_value[i,1], min=0, max=0.0002, log = TRUE) +
    dunif(sig_value[i,2], min=0, max=sig_value[i,1], log = TRUE)

  #-----------------------
  # Use for Bayesian
```

```
  # Approach to Calibration
  #-----------------------

  results[i,]  <- cbind(t(letsgo$x01), t(letsgo$x02),
   sig_value[i,1], sig_value[i,2],p.weight[i])

  #-------------------------
  # Importance Log-Posterior
  # Weights for Resampling
  #-------------------------

  weights[i]   <- p.weight[i] - candidate[i]

}
#-----------------------------------------
# Derive the exponential resampling weights
#-----------------------------------------

weights1    = max(weights) - weights
exp.weights = 1 - (weights1 / sum(weights1))

for (i in 1:1){  # take 1 samples
  x.estimates <- (results[sample(nrow(results),1000,
   replace=TRUE,prob=exp.weights),])
}

Acceptance = length(unique(x.estimates[,2]))/1000


write.table(x.estimates, file = "/home/riversdl/Outputs/results(1_1a).txt",
quote = FALSE, col.names = F, append = T)


#---------------------------
# Rescale the posterior values
#---------------------------

x0.est <- read.table("/home/riversdl/Outputs/results(1_1a).txt", quote="\"")

x0.est <- x0.est[,2:2001]

a = as.numeric(letsgo$a)
```

```
x.00 <- (1/a) * x0.est + mean(X[,2])

#---------------------------
#   Only use this section for
#   the Deterministic Approach
#---------------------------

x01   = as.vector(sapply(x.00[,1:1000],median))
cred1 = sapply(x.00[,1:1000], quantile, probs = c(alpha, 1-alpha))

#---------------------------
#   Only use this section for
#   the Bayesian Approach
#---------------------------

x02   = as.vector(sapply(x.00[,1001:2000],median))
cred2 = sapply(x.00[,1001:2000], quantile, probs = c(alpha, 1-alpha))

#---------------------------
#   Plot time series of estimates
#    with error bars
#---------------------------

# par(mfrow = c(1,1))
#
# t = 2:1000
# errbar(t, x01[2:1000], cred1[2,2:1000], cred1[1,2:1000] ,
# ylab = expression(hat(x)[0[t]]))
# lines(x01[2:1000], col = "blue")                    #Estimate of measurand
# lines(true.x0[2:1000], col = 'dark green', lwd = 2)#True measurand
#
# # par(oma=c(2,2,2,2), mar=c(5.1,4.1,4.1,2.1))
# errbar(t, x02[2:1000], cred2[2,2:1000], cred2[1,2:1000],
# ylab = expression(hat(x)[0[t]]))
# lines(x02[2:1000], col = "purple")                  #Estimate of measurand
# lines(true.x0[2:1000], col = 'red', lwd = 2)        #True measurand

#---------------------------------------
# Calculate MSE and Coverage Probability
#---------------------------------------
burn.in = 5
```

```
pred1 <- matrix(0, (length(true.x0)-burn.in))  #Deterministic
pred2 <- matrix(0, (length(true.x0)-burn.in))  #Bayesian
pred3 <- matrix(0, (length(true.x0)-burn.in))  #Classical
pred4 <- matrix(0, (length(true.x0)-burn.in))  #Inverse
pred5 <- matrix(0, (length(true.x0)-burn.in))  #Hoadley
pred6 <- matrix(0, (length(true.x0)-burn.in))  #Hunter & Lamboy

for(s in 1:(length(true.x0)-burn.in)){

  pred1[s] = (x01[s+burn.in] - true.x0[s+burn.in])**2
  pred2[s] = (x02[s+burn.in] - true.x0[s+burn.in])**2
  pred3[s] = (class$x.pre[s+burn.in] - true.x0[s+burn.in])**2
  pred4[s] = (inver$x.pre[s+burn.in] - true.x0[s+burn.in])**2
  pred5[s] = (hoadl$x.pre[s+burn.in] - true.x0[s+burn.in])**2
  pred6[s] = (huntl$x.pre[s+burn.in] - true.x0[s+burn.in])**2
}

#--------------
# Calculate MSE
#--------------

MSE1      = sum(pred1)/length(pred1)  #Deterministic
MSE2      = sum(pred2)/length(pred2)  #Bayesian
MSE3      = sum(pred3)/length(pred3)  #Classical
MSE4      = sum(pred4)/length(pred4)  #Inverse
MSE5      = sum(pred5)/length(pred5)  #Hoadley
MSE6      = sum(pred6)/length(pred6)  #Hunter & Lamboy

#------------------------------
# Calculate Coverage Probability
#------------------------------

Coverage1 = mean(cred1[1,burn.in:N] < true.x0[burn.in:N] &
cred1[2,burn.in:N] > true.x0[burn.in:N])  #Deterministic
Coverage2 = mean(cred2[1,burn.in:N] < true.x0[burn.in:N] &
cred2[2,burn.in:N] > true.x0[burn.in:N])  #Bayesian
Coverage3 = mean(class$lim[burn.in:N,1] < true.x0[burn.in:N] &
class$lim[burn.in:N,2] > true.x0[burn.in:N])            #Classical
Coverage4 = mean(inver$lim[burn.in:N,1] < true.x0[burn.in:N] &
inver$lim[burn.in:N,2] > true.x0[burn.in:N])            #Inverse
Coverage5 = mean(hoadl$lim[burn.in:N,1] < true.x0[burn.in:N] &
hoadl$lim[burn.in:N,2] > true.x0[burn.in:N])            #Hoadley
```

```
Coverage6 = mean(huntl$lim[burn.in:N,1] < true.x0[burn.in:N] &
huntl$lim[burn.in:N,2] > true.x0[burn.in:N])             #Hunter & Lamboy

#------------------------------
# Calculate Interval Width
#------------------------------

Width1 = mean(cred1[2,burn.in:N] - cred1[1,burn.in:N]) #Deterministic
Width2 = mean(cred2[2,burn.in:N] - cred2[1,burn.in:N]) #Bayesian
Width3 = mean(class$lim[burn.in:N,2] -
              class$lim[burn.in:N,1]) #Classical
Width4 = mean(inver$lim[burn.in:N,2] -
              inver$lim[burn.in:N,1]) #Inverse
Width5 = mean(hoadl$lim[burn.in:N,2] -
              hoadl$lim[burn.in:N,1]) #Hoadley
Width6 = mean(huntl$lim[burn.in:N,2] -
              huntl$lim[burn.in:N,1]) #Hunter & Lamboy

#------------------------------
# Output: Simulation Results
#------------------------------

sim_res <- c(MSE1,Coverage1,Width1,MSE2,Coverage2,Width2,
MSE3,Coverage3,Width3,MSE4,Coverage4,Width4,
MSE5,Coverage5,Width5,MSE6,Coverage6, Width6)

write.table(t(sim_res), file = "/home/riversdl/Outputs/sim_res_1_1a.txt",
quote = FALSE, col.names = F, append = T)

time = proc.time() - ptm

stuff <- c(c,round(MSE1,2), round(MSE2,2), round(MSE3,2),
round(MSE4,2), round(MSE5,2), round(MSE6,2), (time[1]+time[2]))

print(stuff)

 file.remove(file = "/home/riversdl/Outputs/results(1_1a).txt")

 rm(list=ls())

}
```

B Dynamic Nonlinear Calibration R Code

B.1 Data Generating Code

The data generating R script file:

```
simdata <- function(sig1,sig2){

n1 = 1074                    #Initial Simulated Values
T1 = 20

T2 = 40

T3 = 60
T4 = 90
T5 = 100

X1  <- rbind(T1,T2,T3,T4,T5)   #Design Matrix of Reference Temps
X   <- cbind(1,X1,X1^2)

#sig1 = 0.001
#sig2 = 0.0001
B_cov = sig2 * solve(t(X)%*%X)
B.true <- c(-0.000700,0.018580,-0.000117)         #True Beta values
B.hat  <- rmvnorm(n1, B.true, B_cov)
#-------------------------
# Simple Moving Average
#-------------------------
n2  = 75

b0   <- na.omit(SMA(B.hat[,1], n2))
b1   <- na.omit(SMA(B.hat[,2], n2))
b2   <- na.omit(SMA(B.hat[,3], n2))
Beta <- cbind(b0,b1,b2)
#-----------------------------------
# Standard Form of Voltage (Y) values
#-----------------------------------
err        <- rnorm(dim(Beta)[1],sd = sqrt(sig1))#Noise term
err_matrix <- cbind(err,err,err,err,err)         #Matrix of Noise terms
#-------------------------
# Shocks and Disturbances
# Shock1 is the random shocks
```

```
# Shock2 is the sinusoidal disturbance
#--------------------------

shock1 <- c(rep(1,300),rep(1.5,10),rep(0.5,10),
rep(1,250),rep(1.5,10),rep(0.5,10),rep(1,410))

t1 = seq(1,10,by=0.05)
shock2 <- c(rep(1,219),
               (0.25*sin(t1)+1),
               rep(1,219),
               (0.25*sin(t1)+1),
               rep(1,200))
#---------------------------------------
# Vertex Form (h,k) of Voltage (Y) values
#---------------------------------------
h   <- matrix(0,nrow=length(b0))
k   <- matrix(0,nrow=length(b0))
Yt  <- matrix(0,dim(X)[1],length(b0))

a = b2
h = -b1/(2*b2)                            # X-coord of VERTEX
k = b0 - (b1**2)/(4*b2) #+ 0.03*goft[,2] # Y-coord of VERTEX

for(i in 1:length(h)){

Yt[,i] =  (a[i] * shock2[i]) * (X[,2] - h[i])**2 + k[i] + t(err_matrix[i,])

    }
    data <- list(X=X, Yt=Yt)
return(data)
}
```

B.2 Dynamic Nonlinear Calibration Code

The Nonlinear Dynamic Calibration R script file:

```
dlm.Est <- function(X, Y, m.0, C.0, W.t, E.t, y0){

    Xtr = X[,2]
    tot = dim(X)[1]                                          #Num of references
    Sxx = t(X[,2] - mean(X[,2])) %*% (X[,2] - mean(X[,2]))#RSS
    a   = sqrt(tot/Sxx)                                      #Scaling
```

```
  X   = a*(X[,2] - mean(X[,2]))
  X   = matrix(X,tot,1)                                  #Scaled
  X <- cbind(1,X,X**2)

Y <- na.omit(Y)
n <- dim(Y)[1]

R.t <- array(0,c(dim(m.0)[1],dim(m.0)[1],n))#R.t = C.t + W.t
a.t <- matrix(0,dim(m.0)[1],n)              #a.t = m.(t-1)
f.t <- matrix(0,dim(X)[1],n)                #f.t = t(X.t) * a.t
Q.t <- array(0,c(dim(X)[1],dim(X)[1],n))    #Q.t = t(X.t)*R.t*X.t+S.(t-1)
m.t <- matrix(0,dim(m.0)[1],n)              #m.t = a.t + A.t * e.t
C.t <- array(0,c(dim(m.0)[1],dim(m.0)[1],n))#C.t = (R.t * E.t)/Q.t
A.t <- array(0,c(dim(m.0)[1],dim(X)[1],n))  #A.t = (R.t1 * X.t)/Q.t
e.t <- matrix(0,dim(X)[1],n)                #e.t = Y.t - f.t

known.sigma.sq <- matrix(0,n)
post.mean      <- matrix(0,n)
post.var       <- matrix(0,n)

nu1   <- matrix(0,n)
x0    <- matrix(0,n)

SSy <- matrix(0,n)
SSb <- matrix(0,n)
###################################
####  Initialize the vectors  #####
###################################

R.t[,,1] <- C.0 + W.t
a.t[,1]  <- m.0
f.t[,1]  <- X %*% a.t[,1]
Q.t[,,1] <- X %*% R.t[,,1] %*% t(X) + E.t
A.t[,,1] <- (R.t[,,1] %*% t(X)) %*% solve(Q.t[,,1])
e.t[,1]  <- t(Y[1,]) - f.t[,1]
m.t[,1]  <- a.t[,1] + (A.t[,,1] %*% e.t[,1])
C.t[,,1] <- R.t[,,1] - A.t[,,1] %*% Q.t[,,1] %*% t(A.t[,,1])

SSb[1]   <- t(m.t[,1] - m.0) %*% (m.t[,1] - m.0)
SSy[1]   <- t(e.t[,1]) %*% e.t[,1]

#--------------------------
```

```
  # Logic Test of DISCRIMINANT
  #--------------------------

     if(m.t[2,1]**2 < 4 * m.t[3,1] * (m.t[1,1]-y0[1])){

       part11 = 4 * m.t[3,1] * (m.t[1,1]-y0[1])
       part22 = m.t[2,1]**2
     }else{
      part11 = m.t[2,1]**2
      part22 = 4 * m.t[3,1] * (m.t[1,1]-y0[1])
     }

  ################################
  ##  Calibration Distribution  ##
  ################################

#------------------
# Estimate of X0[1]
#------------------

  nu1[1]  = (-m.t[2,1] + sqrt(part11 - part22))/(2 * m.t[3,1])

#----------------------------
# Bayesian Dynamic Calibration
#----------------------------
  mu0     = 0
  tau.sq0 = 1
  n0      = 1
  known.sigma.sq[1] = tr(Q.t[,,1])

  post.mean[1] <- (mu0 * known.sigma.sq[1] + tau.sq0 * nu1[1])/
   (n0 * tau.sq0 + known.sigma.sq[1])
  post.var[1]  <- (known.sigma.sq[1] * tau.sq0)/
                  (n0 * tau.sq0 + known.sigma.sq[1])

#---------------------------------
# Transform estimates back to scale
#---------------------------------

  x0[1]     <- (1/a)*rnorm(1,post.mean[1], sqrt(post.var[1])) + mean(Xtr)

  ################################
```

```
  ####  Bayesian Updating    ####
  ###############################

  for(i in 2:n){

    R.t[,,i] <- C.t[,,i-1] + W.t
    a.t[,i]  <- m.t[,i-1]
    f.t[,i]  <- X %*% a.t[,i]
    Q.t[,,i] <- X %*%  R.t[,,i] %*% t(X) + E.t
    A.t[,,i] <- (R.t[,,i] %*% t(X)) %*% solve(Q.t[,,i])
    e.t[,i]  <- t(Y[i,]) - f.t[,i]
    m.t[,i]  <- a.t[,i] + (A.t[,,i] %*% e.t[,i])
    C.t[,,i] <- R.t[,,i] - A.t[,,i] %*% Q.t[,,i] %*% t(A.t[,,i])

    SSb[i]   <- t(m.t[,i] - m.t[,i-1]) %*% (m.t[,i] - m.t[,i-1])
    SSy[i]   <- t(e.t[,i]) %*% e.t[,i]

    #---------------------------
    # Logic Test of DISCRIMINANT
    #---------------------------

     if(m.t[2,i]**2 < 4 * m.t[3,i] * (m.t[1,i]-y0[i])){

       part11 = 4 * m.t[3,i] * (m.t[1,i]-y0[i])
       part22 = m.t[2,i]**2
     }else{
      part11 = m.t[2,i]**2
      part22 = 4 * m.t[3,i] * (m.t[1,i]-y0[i])
     }

    ###############################
    ## Calibration at point y0 ##
    ###############################

#--------------------
# Estimate of X0[2:n]
#--------------------

   nu1[i]    = (-m.t[2,i] + sqrt(part11 - part22))/(2 * m.t[3,i])

#----------------------------
# Bayesian Dynamic Calibration
```

```
#----------------------------

    known.sigma.sq[i] = tr(Q.t[,,i])

    post.mean[i] <- (mu0 * known.sigma.sq[i] + tau.sq0 * nu1[i])/
     (n0 * tau.sq0 + known.sigma.sq[i])
    post.var[i]  <- abs((known.sigma.sq[i] * tau.sq0)/
     (n0 * tau.sq0 + known.sigma.sq[i]))

    #---------------------------
    # Test for Nan's in Varinace
    #---------------------------
    # ifelse(sqrt(post.var[i])>0, post.var[i],0)

    x0[i]     <- (1/a)*rnorm(1,post.mean[i], sqrt(post.var[i]))+mean(Xtr)
  }

#---------------------------------
# Transform estimates back to scale
#---------------------------------

  nu    <- (1/a)*as.numeric(nu1)+mean(Xtr)

  ####################
  ## Sum of Squares ##
  ####################

  tb1 <-  sum(SSb[1:n])/n
  syy <-  sum(SSy[1:n])/n

  param <- list(x0=x0, tb1=tb1, SSy=syy)

return(param)

}
```

C Dynamic Multi-univariate Linear Calibration R Code

C.1 Data Generating Code

The data generating R script file:

```
simdata <- function(sig1,sig2){  #Function that calls Generated Data

#---------------------
# Generate Sample Data
#---------------------
n1    = 100
n3    = 149
n4    = 50

Xmat  = matrix(c(41,104),ncol=1)
X3    = cbind(1,Xmat)
SigW  = sig2*solve(t(X3)%*%X3)  #[0.00005, 0.001]

Beta1 = rmvnorm(n3,c(1189, 9.153),SigW) #Response of system for device 1
Beta2 = rmvnorm(n3,c(1166, 9.441),SigW) #Response of system for device 2
Beta3 = rmvnorm(n3,c(1199, 9.012),SigW) #Response of system for device 3
B01   = na.omit(SMA(Beta1[,1], n4))
B11   = na.omit(SMA(Beta1[,2], n4))
B02   = na.omit(SMA(Beta2[,1], n4))
B12   = na.omit(SMA(Beta2[,2], n4))
B03   = na.omit(SMA(Beta3[,1], n4))
B13   = na.omit(SMA(Beta3[,2], n4))

SigE  = sig1  #[0.00001, 0.001]

Y3    <- array(0,c(6,1,n1))
Beta3 <- array(0,c(2,3,n1))

#-------------------------------------------
# Generate Stepped and Sinusoidal Disturbances
#-------------------------------------------

# shock1 <- c(rep(1,30),rep(1.05,1),rep(0.95,1),rep(1,25),
rep(1.05,1),rep(0.95,1),rep(1,41))

# t=seq(1,9,by=0.5)
# shock2 <- c(rep(1,27),(0.05*sin(t)+1),rep(1,19),(0.05*sin(t)+1),rep(1,20))

#-------------------------------------------
# Generate Data and Observation Matrices
#-------------------------------------------
```

```
for(h in 1:dim(Y3)[3]){
  Beta3[1,1,h] <- B01[h]
  Beta3[1,2,h] <- B02[h]
  Beta3[1,3,h] <- B03[h]
  Beta3[2,1,h] <- B11[h]
  Beta3[2,2,h] <- B12[h]
  Beta3[2,3,h] <- B13[h]       #Fluctuations (+ shock1[h]) or (+ shock2[h])

  Y3[,,h] <- as.vector(X3 %*% Beta3[,,h] + t(rmvnorm(3,c(0,0),diag(SigE,2))))
  }
data <- list(X3=X3,Y3=Y3)
return(data)
}
```

C.2 Dynamic Multi-univariate Calibration Code

The Dynamic Multi-univariate Calibration `R` script file:

```
dlm.Est <- function(X, Y, W.t, E.t, y0){


  ######################################################
  ####  Prior Distribution for (Theta_0|Data_0)  #####
  ######################################################

  m.0   = matrix(rep(0,dim(Y)[1]),nrow=dim(Y)[1])
  C.0   = diag(100,dim(Y)[1])
  W.t   = diag(W.t,dim(Y)[1])
  E.t   = diag(E.t,dim(Y)[1])

  ######################################################
  ####  Required matrices, Vectors and Arrays  #####
  ######################################################

  dd  = length(Y[,,1])/2                    #Dimension of Identity Matrix

  Xtr = X[1:2,2]
  tot = length(Xtr)                         #Total number of references
  Sxx = t(Xtr-mean(Xtr))%*%(Xtr-mean(Xtr))  #Reference sum of squares
  a   = sqrt(tot/Sxx)                       #Scaling coefficent
  X1  = a*(Xtr - mean(Xtr))                 #Scaled and Centered X
  X   = matrix(X1,tot,1)                    #Scaled reference measurements
```

```
X   = cbind(1,X)
X   = kronecker(diag(dd),X)               #Kronecker Product of X and I

n <- dim(Y)[3]

R.t <- array(0,c(dim(m.0)[1],dim(m.0)[1],n))  #R.t=C.t + W.t
a.t <- matrix(0,dim(m.0)[1],n)                #a.t=m.(t-1)
f.t <- matrix(0,dim(X)[1],n)                  #f.t=t(X.t)*a.t
Q.t <- array(0,c(dim(X)[1],dim(X)[1],n))      #Q.t=t(X.t)*R.t*X.t+E.t
m.t <- matrix(0,dim(m.0)[1],n)                #m.t=a.t+A.t*e.t
C.t <- array(0,c(dim(m.0)[1],dim(m.0)[1],n))  #C.t=(R.t*E.t)/Q.t
A.t <- array(0,c(dim(m.0)[1],dim(X)[1],n))    #A.t=(R.t1*X.t)/Q.t
e.t <- matrix(0,dim(X)[1],n)                  #e.t=Y.t-f.t

B.t <- matrix(0,n,0.5*dim(X)[1])                      #(1xj)Brown(1982)
S.t <- array(0,c(0.5*dim(m.0)[1],0.5*dim(m.0)[1],n))  #(jxj)Brown(1982)

Ymean <- matrix(0,(dim(X)[1]),n) #(qx1) Matrix of Regression Intercepts

known.sigma.sq <- matrix(0,n)
post.mean      <- matrix(0,n)
post.var       <- matrix(0,n)

x.est    <- matrix(0,n)
x01      <- matrix(0,n)
x01.scale <- matrix(0,n)
x.est1   <- matrix(0,n)

SSy <- matrix(0,n)
SSb <- matrix(0,n)

###################################
####  Initialize the vectors  #####
###################################
Ymean[,1]    <- matrix(c(rep(mean(Y[1:2,1,1]),2),
rep(mean(Y[3:4,1,1]),2),rep(mean(Y[5:6,1,1]),2)),ncol=1)

R.t[,,1] <- C.0 + W.t
a.t[,1]  <- m.0
f.t[,1]  <- X %*% a.t[,1]
Q.t[,,1] <- X %*% R.t[,,1] %*% t(X) + E.t
A.t[,,1] <- (R.t[,,1] %*% t(X)) %*% solve(Q.t[,,1])
```

```
e.t[,1]  <- Y[,,1] - Ymean[,1] - f.t[,1]
m.t[,1]  <- a.t[,1] + (A.t[,,1] %*% e.t[,1])
C.t[,,1] <- R.t[,,1] - A.t[,,1] %*% Q.t[,,1] %*% t(A.t[,,1])

SSb[1]   <- t(m.t[,1] - m.0) %*% (m.t[,1] - m.0)
SSy[1]   <- t(e.t[,1]) %*% e.t[,1]

#------------------------
# Brown Estimate at t = 1
#------------------------
B.t[1,]  <- m.t[c(2,4,6),1]
S.t[,,1] <- 0.05*(t(matrix(Y[,,1],2,3)-as.matrix(rep(1,2))%*%
t(as.matrix(Ymean[c(1,3,5),1]))-as.matrix(X1)%*%B.t[1,])%*%
(matrix(Y[,,1],2,3)-as.matrix(rep(1,2))%*%
t(as.matrix(Ymean[c(1,3,5),1]))-as.matrix(X1)%*%B.t[1,]))+0.0001*diag(dd)

################################
##  Calibration Distribution  ##
################################

mu0     = 0
tau.sq0 = 1
n0      = 1
known.sigma.sq[1] = tr(Q.t[1:2,1:2,1])

x.est[1]  = solve(B.t[1,] %*% solve(S.t[,,1]) %*% as.matrix(B.t[1,]))%*%
B.t[1,] %*% solve(S.t[,,1])%*%(t(y0[1,])-Ymean[c(2,4,6),1])

post.mean[1] <- (mu0 * known.sigma.sq[1] + tau.sq0 * x.est[1])/
(n0 * tau.sq0 + known.sigma.sq[1])

post.var[1]  <- (known.sigma.sq[1] * tau.sq0)/
            (n0 * tau.sq0 + known.sigma.sq[1])
x01[1]       <- rnorm(1,  x.est[1],sqrt(post.var[1]))
x01.scale[1] <- (1/a) * x01[1] + mean(Xtr)
x.est1[1]    <- (1/a) * x.est[1] + mean(Xtr)
################################
####  Bayesian Updating    ####
################################

for(i in 2:n){
```

```
Ymean[,i] <- matrix(c(rep(mean(Y[1:2,1,1:i]),2),
rep(mean(Y[3:4,1,1:i]),2),rep(mean(Y[5:6,1,1:i]),2)),ncol=1)

R.t[,,i] <- C.t[,,i-1] + W.t
a.t[,i]  <- m.t[,i-1]
f.t[,i]  <- X %*% a.t[,i]
Q.t[,,i] <- X %*%  R.t[,,i] %*% t(X) + E.t
A.t[,,i] <- (R.t[,,i] %*% t(X)) %*% solve(Q.t[,,i])
e.t[,i]  <- Y[,,i] - Ymean[,i] - f.t[,i]
m.t[,i]  <- a.t[,i] + (A.t[,,i] %*% e.t[,i])
C.t[,,i] <- R.t[,,i] - A.t[,,i] %*% Q.t[,,i] %*% t(A.t[,,i])

SSb[i]   <- t(m.t[,i] - m.t[,i-1]) %*% (m.t[,i] - m.t[,i-1])
SSy[i]   <- t(e.t[,i]) %*% e.t[,i]

#---------------------------
# Brown Estimate at t = 2...T
#---------------------------

B.t[i,]  <- m.t[c(2,4,6),i]
S.t[,,i] <- 0.05*(t(matrix(Y[,,i],2,3)-as.matrix(rep(1,2))%*%
t(as.matrix(Ymean[c(1,3,5),i]))-as.matrix(X1)%*%B.t[i,])%*%
(matrix(Y[,,i],2,3)-as.matrix(rep(1,2))%*%
t(as.matrix(Ymean[c(1,3,5),i]))-as.matrix(X1)%*%B.t[i,]))+0.0001*diag(dd)

############################
## Calibration at point y0 ##
############################

 mu0     = 0
 tau.sq0 = 1
 n0      = 1
known.sigma.sq[i] = tr(Q.t[1:2,1:2,i])

x.est[i]     <- solve(B.t[i,] %*% solve(S.t[,,i]) %*%
             as.matrix(B.t[i,]))%*% B.t[i,] %*%
             solve(S.t[,,i])%*%(t(y0[i,])-Ymean[c(2,4,6),i])

post.mean[i] <- (mu0 * known.sigma.sq[i] + tau.sq0 * x.est[i])/
(n0 * tau.sq0 + known.sigma.sq[i])

post.var[i]  <- (known.sigma.sq[i] * tau.sq0)/
```

```
                (n0 * tau.sq0 + known.sigma.sq[i])
    x01[i]      <- rnorm(1,x.est[i],sqrt(post.var[i]))
    x01.scale[i] <- (1/a) * x01[i] + mean(Xtr)
    x.est1[i]    <- (1/a) * x.est[i] + mean(Xtr)
  }

  ####################
  ## Sum of Squares ##
  ####################

  sbb <-  sum(SSb[1:n])/n
  syy <-  sum(SSy[1:n])/n

  param <- list(x0=x01.scale, x02=x.est1, meanX0=post.mean,
  sigX0=post.var, SSb=sbb, SSy=syy, a=a)
  return(param)

}
```

D LinCal R Package

Package 'LinCal'

November 8, 2014

Title Static Univariate Frequentist and Bayesian Linear Calibration

Version 1.0

Author Derick L. Rivers <riversdl@vcu.edu> and Edward L. Boone

Maintainer Derick L. Rivers <riversdl@vcu.edu>

Description Estimate and confidence/credible intervals for an unknown regressor x0 given an observed y0.

Depends R (>= 3.0.2)

License GPL-2

LazyData yes

NeedsCompilation no

Repository CRAN

Date/Publication 2014-11-08 02:12:12

R topics documented:

LinCal-package *Static Univariate Frequentist and Bayesian Linear Calibration*

Description

A collection of R functions for conducting linear statistical calibration.

Details

Package:	LinCal
Type:	Package
Version:	1.0
Date:	2014-11-06
License:	GPL-2

Author(s)

Derick L. Rivers and Edward L. Boone

Maintainer: Derick L. Rivers <riversdl@vcu.edu>

References

Eisenhart, C. (1939). The interpretation of certain regression methods and their use in biological and industrial research. Annals of Mathematical Statistics. 10, 162-186.

Krutchkoff, R. G. (1967). Classical and Inverse Regression Methods of Calibration. Technometrics. 9, 425-439.

Hoadley, B. (1970). A Bayesian look at Inverse Linear Regression. Journal of the American Statistical Association. 65, 356-369.

Hunter, W., and Lamboy, W. (1981). A Bayesian Analysis of the Linear Calibration Problem. Technometrics. 3, 323-328.

Examples

```
library(LinCal)

data(wheat)

plot(wheat[,6],wheat[,2])

## Classical Approach
class.calib(wheat[,6],wheat[,2],0.05,105)

## Inverse Approach
```

```
inver.calib(wheat[,6],wheat[,2],0.05,105)

## Bayesian Inverse Approach
hoad.calib(wheat[,6],wheat[,2],0.05,105)

##Bayesian Classical Approach
huntlam.calib(wheat[,6],wheat[,2],0.05,105)
```

`class.calib` *Classical Linear Calibration Function*

Description

`class.calib` uses the classical frequentist approach to estimate an unknown X given observed vector y0 and calculates confidence interval estimates.

Usage

```
class.calib(x, y, alpha, y0)
```

Arguments

`x`	numerical vector of regressor measurments
`y`	numerical vector of observation measurements
`alpha`	the confidence interval to be calculated
`y0`	vector of observed calibration value

References

Eisenhart, C. (1939). The interpretation of certain regression methods and their use in biological and industrial research. Annals of Mathematical Statistics. 10, 162-186.

Examples

```
X <- c(1,1,2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10)
Y <- c(1.8,1.6,3.1,2.6,3.6,3.4,4.9,4.2,6.0,5.9,6.8,6.9,8.2,7.3,8.8,8.5,9.5,9.5,10.6,10.6)

class.calib(X,Y,0.05,6)
```

hoad.calib *Bayesian Inverse Linear Calibration Function*

Description

hoad.calib uses an inverse Bayesian approach to estimate an unknown X given observed vector y0 and calculates credible interval estimates.

Usage

```
hoad.calib(x, y, alpha, y0)
```

Arguments

x	numerical vector of regressor measurments
y	numerical vector of observation measurements
alpha	the confidence interval to be calculated
y0	vector of observed calibration value

References

Hoadley, B. (1970). A Bayesian look at Inverse Linear Regression. Journal of the American Statistical Association. 65, 356-369.

Examples

```
X <- c(1,1,2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10)
Y <- c(1.8,1.6,3.1,2.6,3.6,3.4,4.9,4.2,6.0,5.9,6.8,6.9,8.2,7.3,8.8,8.5,9.5,9.5,10.6,10.6)

hoad.calib(X,Y,0.05,6)
```

huntlam.calib *Bayesian Classical Linear Calibration Function*

Description

huntlam.calib uses the classical Bayesian approach to estimate an unknown X given observed vector y0 and calculates credible interval estimates.

Usage

```
huntlam.calib(x, y, alpha, y0)
```

Arguments

x	numerical vector of regressor measurments
y	numerical vector of observation measurements
alpha	the confidence interval to be calculated
y0	vector of observed calibration value

References

Hunter, W., and Lamboy, W. (1981). A Bayesian Analysis of the Linear Calibration Problem. Technometrics. 3, 323-328.

Examples

```
X <- c(1,1,2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10)
Y <- c(1.8,1.6,3.1,2.6,3.6,3.4,4.9,4.2,6.0,5.9,6.8,6.9,8.2,7.3,8.8,8.5,9.5,9.5,10.6,10.6)

huntlam.calib(X,Y,0.05,6)
```

---

inver.calib *Inverse Linear Calibration Function*

Description

inver.calib uses the inverse frequentist approach to estimate an unknown X given observed vector y0 and calculates confidence interval estimates.

Usage

```
inver.calib(x, y, alpha, y0)
```

Arguments

x	numerical vector of regressor measurments
y	numerical vector of observation measurements
alpha	the confidence interval to be calculated
y0	vector of observed calibration value

References

Krutchkoff, R. G. (1967). Classical and Inverse Regression Methods of Calibration. Technometrics. 9, 425-439.

Examples

```
X <- c(1,1,2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10)
Y <- c(1.8,1.6,3.1,2.6,3.6,3.4,4.9,4.2,6.0,5.9,6.8,6.9,8.2,7.3,8.8,8.5,9.5,9.5,10.6,10.6)

inver.calib(X,Y,0.05,6)
```

wheat *Percentage Water, Percentage Protein, and Infrared Reflectance Measurements of Hard Wheat*

Description

A dataset containing 21 samples of hard wheat. The variables are as follows:

Usage

```
data("wheat")
```

Format

A data frame with 21 observations on the following 6 variables.

- `Y1` infrared reflectance vector
- `Y2` infrared reflectance vector
- `Y3` infrared reflectance vector
- `Y4` infrared reflectance vector
- `X1` percentage water vector
- `X2` percentage protein vector

Source

Brown, P. J. (1982). Multivariate calibration. Journal of the Royal Statistical Society B. 44, 287-321.

Examples

```
data(wheat)
## maybe str(wheat) ; plot(wheat) ...
```

Index

Vita

Derick Lorenzo Rivers was born on June 12, 1970 in Richmond, Virginia. After graduating from George Wythe High School in 1989, he received his Bachelor of Science degree in Mathematics in 1995 from Morris College, Sumter, South Carolina. He received a teaching certificate and taught for several years in Richmond Public School System as a Mathematicsbefore returning to graduate school. He received a Master of Education in Mathematics Education in 2004 from Cambridge College, Cambridge, Massachusetts. Upon completion of his Master of Education, Rivers held the position of Lead Mathematics Instructor for ECPI University (Richmond campuses) from 2004 until 2010. Rivers received numerous accolades during his tenure at ECPI University. He earned the following: "Most Innovative Instructor Award" in 2006; "Most Effective Instructor Award" in 2007; "Most Effective Instructor Award" in 2008; and the "Instructor of the Year Award" in 2010.

During his time at ECPI University, Rivers felt the need to return to the role of student if he wanted to be a better educator. He earned a Master of Science degree in Mathematical Science with a concentration in Statistics from Virginia Commonwealth University in 2009. Being persuaded by faculty at VCU, Rivers entered the newly formed Doctor of Philosophy in Systems Modeling and Analysis program. Taking the suggestion of his advisor Ed Boone, he submitted a proposal for a research grant that would support him during his studies. Derick Rivers was awarded the NASA Harriet G. Jenkins Predoctoral Fellowship which provided him with $121,500 of support from September 2011 until August 2014.

Rivers continued his passion for teaching by facilitating a lab course for STAT 208 Statistical Thinking and several sections of HPEX 334 Measurement and Analysis in Teaching and Exercise Science across three semesters. Currently, he and Ed Boone have a version of Chapter 2 (A Dynamic Approach to Linear Statistical Calibration with an Application in Microwave Radiometry) under review with the *Journal of Statistical Computation and Simulation*. Derick Rivers and Ed Boone also have a version of Chapter 3 (Dynamic Bayesian Nonlinear Calibration) under review with *Computational Statistics and Data Analysis*. Rivers is expecting to submit Chapter 4 to a peer-reviewed journal, though the destination is not known yet. He has a contributed `R` package on CRAN package repository titled **LinCal** which is used to conduct static univariate linear calibration from frequentist and Bayesian perspectives.