2009-11-20

Improving Accuracy in Microwave Radiometry via Probability and Inverse Problem Theory

Derek Lavell Hudson
Brigham Young University - Provo

BYU ScholarsArchive Citation
Hudson, Derek Lavell, "Improving Accuracy in Microwave Radiometry via Probability and Inverse Problem Theory" (2009). *Theses and Dissertations*. 1945.
https://scholarsarchive.byu.edu/etd/1945

IMPROVING ACCURACY IN MICROWAVE RADIOMETRY VIA PROBABILITY AND INVERSE PROBLEM THEORY

by

Derek L. Hudson

A dissertation submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Electrical and Computer Engineering

Brigham Young University

December 2009

ABSTRACT

IMPROVING ACCURACY IN MICROWAVE RADIOMETRY VIA PROBABILITY AND INVERSE PROBLEM THEORY

Derek L. Hudson
Electrical and Computer Engineering
Doctor of Philosophy

Three problems at the forefront of microwave radiometry are solved using probability theory and inverse problem formulations which are heavily based in probability theory. Probability theory is able to capture information about random phenomena, while inverse problem theory processes that information. The use of these theories results in more accurate estimates and assessments of estimate error than is possible with previous, non-probabilistic approaches. The benefits of probabilistic approaches are expounded and demonstrated.

The first problem to be solved is a derivation of the error that remains after using a method which corrects radiometric measurements for polarization rotation. Yueh [1] proposed a method of using the third Stokes parameter T_U to correct brightness temperatures such as T_v and T_h for polarization rotation. This work presents an extended error analysis of Yueh's method. In order to carry out the analysis, a forward model of polarization rotation is developed which accounts for the random nature of thermal radiation, receiver noise, and (to first order) calibration. Analytic formulas are then derived and validated for bias, variance, and root-mean-square error

(RMSE) as functions of scene and radiometer parameters. Examination of the formulas reveals that: 1) natural T_U from planetary surface radiation, of the magnitude expected on Earth at L-band, has a negligible effect on correction for polarization rotation; 2) RMSE is a function of rotation angle Ω, but the value of Ω which minimizes RMSE is not known prior to instrument fabrication; and 3) if residual calibration errors can be sufficiently reduced via postlaunch calibration, then Yueh's method reduces the error incurred by polarization rotation to negligibility.

The second problem addressed in this dissertation is optimal estimation of calibration parameters in microwave radiometers. Algebraic methods for internal calibration of a certain class of polarimetric microwave radiometers are presented by Piepmeier [2]. This dissertation demonstrates that Bayesian estimation of the calibration parameters decreases the RMSE of the estimates by a factor of two as compared with algebraic estimation. This improvement is obtained by using knowledge of the noise structure of the measurements and by utilizing all of the information provided by the measurements. Furthermore, it is demonstrated that much significant information is contained in the covariance information between the calibration parameters. This information can be preserved and conveyed by reporting a multidimensional pdf for the parameters rather than merely the means and variances of those parameters. The proposed method is also extended to estimate several hardware parameters of interest in system calibration.

The final portion of this dissertation demonstrates the advantages of a probabilistic approach in an empirical situation. A recent inverse problem formulation, sketched in [3], is founded on probability theory and is sufficiently general that it can be applied in empirical situations. This dissertation applies that formulation to the retrieval of Antarctic air temperature from satellite measurements of microwave brightness temperature. The new method is contrasted with the curve-fitting approach which is the previous state-of-the-art. The adaptibility of the new method not only results in improved estimation but is also capable of producing useful estimates of air temperature in areas where the previous method fails due to the occurence of melt events.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

"Truth is knowledge of things as they are, and as they were, and as they are to come." –Doctrine and Covenants 93:24

Many problems in microwave radiometry are currently approached and solved in an algebraic fashion. This approach is favored because it is simple and accessible to audiences unfamiliar with more sophisticated approaches. However, the use of probability theory and probabilistic inverse problem theory can yield solutions which are significantly more accurate than algebraic solutions.

The foundation for an accurate solution is an accurate forward model. While a purely algebraic model is easy to handle, a probabilistic model can include additional information. One avenue for this is to model quantities as random variables rather than simple algebraic entities. An additional avenue is to capture the relationships among variables using probability density functions (pdfs) rather than deterministic functions.

A forward model with probabilistic variables and relationships should also be inverted probabilistically in order to obtain information on parameters of interest. Bayes' theorem can be used to perform such an inversion when an algebraic forward model exists. A more general approach, recently been published by Tarantola [3], is available to invert probabilistic forward models.

1.1 Purpose

A primary purpose of this dissertation is to manifest advantages obtained by employing probability theory and probabilistic inverse problem theory instead of simple algebraic methods. These advantages are: (a) more accurate estimates of parameters, (b) more accurate assessments of error in estimates, and (c) conservation of important covariance information.

Applying probability theory and probabilistic inverse problem theory is not necessarily a straight-forward task. Hence, a second purpose of this dissertation is to demonstrate details of application to three problems of increasing difficulty. This prepares the way for others, especially those in the field of microwave radiometry, to more easily apply probability theory and probabilistic inverse problem theory in other problems.

The third purpose of this dissertation is to solve three actual problems at the forefront of microwave radiometry. These problems and the results obtained are now summarized.

1.2 Outline of Problems, Solutions, and Contributions

1.2.1 Analysis of Correction for Polarization Rotation

The first problem to be solved (Chapter 3) is a derivation of the accuracy of a procedure which corrects for unwanted rotation in the polarization basis of radiometric measurements. Yueh [1] proposed a method of using the third Stokes parameter T_U to correct brightness temperatures such as T_v and T_h for polarization rotation. The method is termed "polarization rotation correction" (PRC). Chapter 3 presents an extended error analysis of PRC. In order to carry out the analysis, a forward model of polarization rotation is developed which accounts for the random nature of thermal radiation, receiver noise, and (to first order) calibration. An appendix to Chapter 3 derives a covariance matrix which is used in Chapters 3 and 4. Analytic formulas are then derived for the residual bias, variance, and root-mean-square error (RMSE) of PRC as functions of scene and radiometer parameters. These formulas are validated independently via Monte Carlo simulation.

Examination of the formulas yields valuable insights. In the planning of an L-band spaceborne radiometer for soil moisture sensing, there has been concern that natural T_U from planetary surface radiation will significantly degrade the performance of PRC. However, the formulas derived in Chapter 3 indicate that natural T_U of the magnitude expected on Earth at L-band (less than 5 K) has a negligible effect on PRC for such a sensor.

Another insight provided by the derived formulas is that residual error (after PRC) stems from residual calibration errors. Furthermore, it is more accurate to consider rotation angle Ω to be a modulator of residual calibration errors, rather than an error source itself. In addition, if residual calibration errors are unknown, then the value of Ω which minimizes residual PRC error is also unknown. This contradicts the notion that error automatically increases with Ω (that notion is true when no correction is made). This work is published as [4].

1.2.2 Optimal Internal Calibration

The second problem addressed in this dissertation (Chapter 4) is optimal estimation of calibration parameters in microwave radiometers. Algebraic methods for internal calibration of a certain class of polarimetric microwave radiometers are presented by Piepmeier [2]. Chapter 4 and its appendix work out the intricacies necessary to estimate the calibration parameters probabilistically (via Bayes theorem) rather than algrebraically. The result, validated by Monte Carlo simulation, is a reduction in RMSE by a factor of two as compared with algebraic estimation. This improvement is possible because the probabilistic method is able to utilize redundancy in the measurements (there are 16 measurements and only ten parameters to estimate). Optimal utilization of the redundancy is possible because probability theory captures knowledge of the noise structure of the measurements. While the probabilistic method has greater complexity and requires more computation than an algebraic method, the complexity is worked out in this dissertation and the computation is minor by today's standards.

Chapter 4 also demonstrates that much significant information is available as covariance information between the calibration parameters. This information can be preserved and conveyed by reporting a single multidimensional pdf for the parameters as the solution to the problem, rather than merely reporting the means and variances of the parameters.

As an extension of the work, pdfs are also found which characterize and/or estimate eight hardware parameters of this class of radiometer. One parameter, the amplitude imbalance of a hybrid coupler, receives no uncertainty from the effects modeled in this paper. Another parameter, a bandpass equalization efficiency, is also well resolved. The remaining six parameters cannot be individually resolved from radiometer measurements. However, either ratios or products of pairs of them can be well resolved.

This work is published as [5]. Note that the theoretical advantages predicted by this dissertation could be optimistic. Application of the probablistic method in hardware is needed in order to validate the predictions herein.

1.2.3 Improved Retrieval of Polar Air Temperature

The third problem to be solved in this dissertation (Chapter 5) demonstrates the advantages of a probabilistic approach in an empirical situation. A recent inverse problem formulation, sketched in [3], is strongly based in probability theory and is sufficiently general that it can be applied in empirical situations. Chapter 5 applies that formulation to the retrieval of Antarctic air temperature from satellite measurements of microwave brightness temperature.

The previous state-of-the-art method (termed "sinusoidal method") used a sinusoidal curve to approximate the relationship between Antarctic air temperature T and satellite measurements of microwave brightness temperature B. The new method (termed "pdf method") achieves better performance because it derives a more exact (albeit empirical) relationship between T and B. The pdf method is more accurate because it does not constrain the relationship between T and B to a sinusoidal model. Instead, coincident T and B data are treated as samples of an underlying two-dimensional pdf which relates T and B.

To compare the sinusoidal and pdf methods, one or two years of coincident T and B data at an Antarctic weather station are used to train both methods (that is, to derive an empirical relationship between T and B at the station) . The methods are then tested on another year of B data. The output of each method ($\hat{T}$) is then compared with station T data. In particular, the root-mean-square (RMS) of the difference is calculated.

This procedure demonstrates that the adaptibility of the new method results in improved estimation. With only one year of training data, the performance of the pdf method is slightly better than the sinusoidal method. This holds true over a wide range of two controlling parameters, σ_D and σ. These parameters are standard deviations which specify the credibility which the pdf method gives at a particular site to its two sources of information (namely, day of year and 37-GHz v-pol brightness temperature measured on that day).

When two years of training data are available, the pdf method can be self-trained to make intelligent choices for σ and σ_D. This intelligence or adaptivity produces significant improvements. Compared to the sinusoidal method, the pdf method reduces RMS error by an average of 0.3 °C at the three inland sites with sufficient data for use in this study.

With two years of training data, the adaptibility of the new method also makes it capable of producing useful estimates of air temperature in areas where the previous method fails due to the occurence of melt events. Two sites have sufficient data for this study. Both are located on the Antarctic coast. Compared to the sinusoidal method, the pdf method reduces RMS error by an average of 8.4 °C at these sites.

A final advantage of the pdf method is that the variance of a T estimate can be estimated even when ground truth data are unavailable. Specifically, the variance can be calculated directly from the marginal pdf for T (this marginal is an output of the pdf method).

1.3 Dissertation Structure

The dissertation is developed as follows. Chapter 2 provides background concepts on microwave polarimetric radiometry and estimation theory. Chapter 3 is an error analysis of polarization rotation correction in microwave radiometry. Chapter 4 derives optimal (Bayesian) internal calibration of microwave radiometers. Chapter 5 derives an improved method of estimating polar air temperature from microwave radiometric data. Chapter 6 contains summary and discussion, including future work options. The appendices include a number of detailed derivations and a description of the data sets used in Chapter 5.

Chapter 2

Background

The three problems solved in this dissertation all concern microwave radiometry and estimation theory. A few essential concepts from these fields are summarized here. More complete expositions of radiometry are found in [6, Chs. 4 and 6] and [7, Ch. 7]. A good introduction to classical probability and estimation theory is [8], with deeper developments in [9] and [3]. Background information which relates to only one of the specific problems in this dissertation is presented within the relevant chapter.

2.1 Microwave Radiometry

Microwave radiometry is the measurement of natural electromagnetic radiation in the microwave spectrum. This radiation is generated by the tiny vibrations, collisions, and other movements which constitute the thermal energy of matter. The characteristics of this radiation are described by Planck's radiation law and additional theory.

The intensity of this radiation is measured by a radiometer as the brightness temperature T_B of a scene at a certain frequency and polarization. This brightness is the product of the physical temperature T of the scene and a quantity called the emissivity ϵ of the scene:

$$T_B = \epsilon T. \tag{2.1}$$

The emissivity ranges from 0 to 1. It depends on the polarization of the radiation and on the composition, orientation, and texture of the scene. Both T_B and T are measured in kelvins, while ϵ is unitless.

A phenomenon that is central to all three problems in this dissertation is the random fluctuation of T_B measurements. Even when ϵ and T are constant, the randomness of the tiny movements of the atoms and molecules in a scene causes the intensity of the emitted radiation to fluctuate about a mean value. The mean value is the quantity which we wish to measure. The fluctuation is modeled by a zero mean random variable which is termed noise equivalent ΔT ($NE\Delta T$):

$$\text{measured } T_B \equiv \hat{T}_B = T_B + \Delta T. \tag{2.2}$$

An important characteristic of $NE\Delta T$ is that its standard deviation (STD) is equal to $T_B/\sqrt{B\tau}$, where B is the bandwidth of the radiation measured by the radiometer and τ is the length of time over which the radiation is measured. Also, ΔT can be approximated as a Gaussian random variable for $B\tau > 10$ [7, Ch. 7].

The basic process of measuring T_B is straightforward. An antenna collects the microwave radiation and converts it to a small electrical signal. This signal is amplified and filtered by a receiver. The signal power, which is proportional to T_B, is then measured ("detected") and digitized. This process can be represented by

$$P = G\left(T_B + T_{RX}\right), \tag{2.3}$$

where G is the overall amplification (gain) and T_{RX} is additive noise power produced by the receiver electronics. The radiometer consists of all the hardware to measure T_B from the antenna through the digitizing device. The calibration of the radiometer consists of obtaining estimates of G and T_{RX}, using targets or inputs with known brightness temperature. Using hat marks to denote estimates, the desired estimate of T_B is then

$$\hat{T}_B = \frac{P}{\hat{G}} - \hat{T}_{RX}. \tag{2.4}$$

2.2 Polarization and Stokes Parameters

The polarization of microwave radiation is a characteristic of high significance to many users of radiometer data. This section sketches the basic principles of polarization characterization.

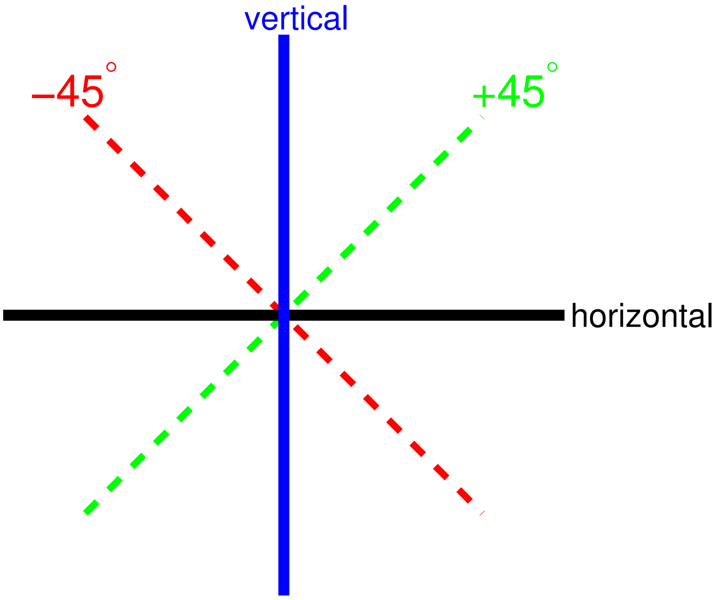

Figure 2.1: Axes for defining the polarization of light. The plane of the figure is perpendicular to the direction of travel of the light.

Consider the plane perpendicular to the direction of travel of a light wave (such as microwave radiation). This plane is depicted in Figure 2.1. The electric field component of light oscillates in this plane. If the oscillation is back and forth along the vertical axis, then the wave is said to be vertically polarized. Likewise, a horizontally polarized wave is one whose electric field oscillations occur in the direction of the horizontal axis[1]. Two other polarizations are those at plus and minus 45° to the vertical and horizontal axes, shown by dashed lines in Figure 2.1. Finally, a circularly polarized wave is one in which the axis of electric field oscillations rotates smoothly and continually rather than remaining fixed.

---

[1] For remote sensing of the earth, horizontal is defined as the axis parallel to both the earth's surface and the plane in Figure 2.1. Vertical is defined as the remaining axis (perpendicular to both the direction of travel and the horizontal axis).

The radiation (light) measured by a radiometer is the sum of large numbers of tiny waves which are randomly emitted by different parts of the scene. The instantaneous axis of oscillation of the total electric field is therefore rather random – that is, unpolarized to some degree. Nevertheless, in most cases the radiation is partially polarized – that is, the oscillations have a larger component along one axis than along others.

The degree to which such "random" radiation is polarized in various manners can be described as follows. The light can be filtered such that only the vertically polarized component of the oscillations is preserved. The brightness temperature of this filtered radiation is defined as T_v. Similarly, brightness temperatures of the radiation filtered along the other axes in Figure 2.1 are denoted by T_h, T_{+45}, and T_{-45}. It is also possible to measure the brightness temperature of only the circularly polarized components of the radiation. The component whose axis of oscillation rotates to the right is denoted T_r while the component whose axis of oscillation rotates to the left is denoted T_l.

The polarization state of random radiation can be completely summarized by four parameters known as Stokes parameters after their inventor. These are

$$\begin{aligned} T_I \text{ (``first Stokes parameter'')} &\equiv T_v + T_h = T_{+45} + T_{-45} = T_l + T_r, && (2.5)\\ T_Q \text{ (``second Stokes parameter'')} &\equiv T_v - T_h, && (2.6)\\ T_U \text{ (``third Stokes parameter'')} &\equiv T_{+45} - T_{-45}, && (2.7)\\ T_V \text{ (``fourth Stokes parameter'')} &\equiv T_l - T_r. && (2.8) \end{aligned}$$

It is also common to denote the third and fourth Stokes parameters by T_3 and T_4, respectively. Radiation which is completely unpolarized has nonzero T_I while $T_Q = T_U = T_V = 0$. More generally, the degree of polarization of radiation is given by $p = \frac{\sqrt{T_Q^2+T_U^2+T_V^2}}{T_I}$, with p ranging from zero to one.

2.3 Estimation Theory

This section provides a brief sketch of key selected concepts from estimation theory which are frequently referred to throughout the dissertation. Specifically, it touches on metrics for error and on the concept of forward versus inverse problems.

2.3.1 Error Statistics

The error in an estimator $\hat{Q}$ of a quantity with true value Q can be defined as follows. If the same process (estimator) is used to estimate the quantity N times, under identical conditions every time, then an ensemble of estimates is obtained[2], denoted $\hat{Q}_i$ with $i = 1, ..., N$. The root-mean-square error (RMSE) in this set of estimates is an average error defined by

$$RMSE(\text{set of } \hat{Q}_i) \equiv \sqrt{\frac{\left(Q - \hat{Q}_1\right)^2 + ... + \left(Q - \hat{Q}_N\right)^2}{N}}. \tag{2.9}$$

As N grows to infinity, the arithmetic mean (average) of the ensemble is equivalent (for practical purposes and the problem at hand) to the “expected value” of the ensemble and is denoted $< \cdot >$. The expected RMSE of the estimator $\hat{Q}$ is defined as

$$RMSE(\hat{Q}) \equiv \sqrt{< (Q - \hat{Q})^2 >}. \tag{2.10}$$

It is often very natural and useful to decompose RMSE into the root-sum-square (RSS) of a bias and a standard deviation (STD). Bias is the expected difference between the true Q and the $\hat{Q}_i$, while the STD captures the fluctuation of the $\hat{Q}_i$ around their mean value. Defining this mean value with $\bar{Q} \equiv < \hat{Q} >$, then

$$\begin{aligned} Bias(\hat{Q}) &\equiv < Q - \hat{Q} > \quad = Q - \bar{Q}, && (2.11)\\ STD(\hat{Q}) &\equiv \sqrt{< (\bar{Q} - \hat{Q})^2 >}, && (2.12)\\ RMSE(\hat{Q}) &= \sqrt{Bias^2(\hat{Q}) + STD^2(\hat{Q})}. && (2.13) \end{aligned}$$

---

[2] Typically it is not possible to physically obtain an ensemble since conditions do not remain constant. Thus, the idea of an ensemble is primarily theoretical.

2.3.2 Forward and Inverse Problems

A *forward* model or problem seeks to show how measurements M are a function of parameters Q which characterize a particular situation. Conversely, an *inverse* model or problem seeks to determine Q based on measurements M.

One type of forward model is a probability density function (pdf) for M, given a known set of Q. This pdf is denoted $p(M|Q)$. In many cases it is possible to find a solution to the inverse problem as another pdf, $p(Q|M)$, by using the following form of **Bayes' theorem**:

$$p(Q|M) = \frac{p(M|Q)p(Q)}{p(M)}. \tag{2.14}$$

Here, $p(Q)$ is information on Q from some source other than the measurements M (prior information on Q). Chapter 4 provides an illustration of the use of (2.14). The denominator is simply a constant, a normalizing factor which creates a valid pdf for fixed M.[3]

---

[3] This constant is left undetermined for the problems addressed in this dissertation. If it is needed, it can be found as

$$p(M) = \int p(M|Q)p(Q)dQ. \tag{2.15}$$

Chapter 3

A Probabilistic Analysis of Polarization Rotation Correction

This chapter illustrates the use of probability theory to analyze the residual error in a correction technique. Previously, this residual error was modeled and analyzed in a simple algebraic fashion [1].

While the algebraic approach is straightforward and accessible, the residual error under examination is known to contain random constituents. Therefore, this chapter expands the error model to include random variables. Analysis of this probabilistic error model is significantly more challenging than for the previous algebraic model. However, it yields more accurate characterization of the residual error. This analysis also yields valuable insight into the sources of residual error. This work is published as [4]. It demonstrates how error analysis in microwave radiometry can be taken to a higher level of fidelity.

3.1 Introduction

The earth's ionosphere and magnetic field cause Faraday rotation of the polarization of radiation emanating from the earth's surface. This rotation mixes the vertical and horizontal polarization components of brightness temperatures, T_v and T_h, degrading the measurement of both. The oft-used second Stokes parameter, T_Q ($\equiv T_v - T_h$), is doubly degraded. For L-band satellite measurements, the error in T_Q due to uncorrected Faraday rotation can exceed 10 K, depending on solar activity, incidence angle, and the angle between the look direction and the Earth's magnetic field [10]. (Faraday rotation is inversely proportional to the square of frequency. Therefore this source of polarization rotation is less important above L-band.)

Additional polarization rotation occurs if a sensor's antenna feed polarization basis is rotated with respect to the natural polarization basis of the earth's surface. Such rotation may occur as an accidental misalignment [11] or may be deliberately permitted in order to simplify hardware [12].

Near-future L-band spaceborne radiometers, namely SMOS [13] and Aquarius [14], are being designed to perform polarization rotation correction (PRC) in post-processing. A basic method involves measuring the third Stokes parameter, T_U, in addition to the usual T_v and T_h. The method is introduced by Yueh in [1].

Previously developed forward models of polarization rotation [15], [12], [1], [11], are deterministic and neglect the role of receiver channel noise (although [16] includes noise in simulations). In Appendix A, an extended model is developed which takes into account the random nature of the radiation and also accounts for receiver noise. Simple and accurate expressions are derived for the means, variances, and covariances of the measurements in a three-channel (T_v, T_h, and T_U) radiometer. These are derived in Appendix A and summarized in Section 3.2.

In Section 3.3, Yueh's correction technique is reviewed. In Section 3.4 derivations are given for the mean, variance, and root mean squared error (RMSE) of the resulting estimate of T_Q. Similar derivations for T_v and T_h are presented in Section 3.5. Insights from the resulting formulas are presented in Section 3.6, and conclusions are offered in Section 5.6.

Throughout this work, case studies are shown for the Aquarius radiometer, whose deployment is expected in 2010. The Aquarius instrument will have three beams with respective incidence angles of $28.7°, 37.8°$, and $45.6°$ [14]. For measurements of ocean emissions, these angles dictate a nominal T_Q of about 20, 35, and 53 K, respectively [17]. The Aquarius instrument also has nominal integration time τ of 6 seconds. Reference is also made to the canceled NASA Hydros mission [18], whose nominal τ was 0.016 seconds. The Hydros incidence angle is adjusted from 39.3° to 37.8° in this study, to match Aquarius.

3.2 Summary of Forward Model

As shown in Appendix A, the processes of receiving, detecting, and calibrating the first three Stokes parameters in a polarimetric radiometer can be summarized with the forward model

$$\hat{T}_{Ia} = T_I + \Delta T_{RX,I} + \Delta T_{sys,I}, \tag{3.1}$$

$$\hat{T}_{Qa} = +T_Q \cos 2\Omega + T_U \sin 2\Omega + \Delta T_{RX,Q} + \Delta T_{sys,Q}, \tag{3.2}$$

$$\hat{T}_{Ua} = -T_Q \sin 2\Omega + T_U \cos 2\Omega + \Delta T_{RX,U} + \Delta T_{sys,U}. \tag{3.3}$$

The quantities on the left hand sides are our measurements of the first three Stokes parameters after rotation, detection, and calibration. On the right sides, the natural Stokes parameters of the scene, T_I, T_Q, and T_U, are altered by polarization rotation Ω [15] and perturbed by error sources, represented by quantities with a Δ prefix.

$\Delta T_{RX,I}$, $\Delta T_{RX,Q}$, and $\Delta T_{RX,U}$ are residual biases from the calibration process that is performed throughout data collection. For example, if the fourth calibration scheme described in [2] is used, then $\Delta T_{RX,U}$ corresponds to all but the first term on the right side of (42) in [2]. That same calibration scheme also leads to

$$\Delta T_{RX,v} = \Delta T_{RX,h} = \frac{T_H \hat{T}_C - T_C \hat{T}_H}{T_H - T_C}, \tag{3.4}$$

where T_H and T_C are the true temperatures of the hot and cold calibration sources, while $\hat{T}_H$ and $\hat{T}_C$ are the best available estimates of them. (The calibration sources could be noise diodes or external targets, for example.) Then $\Delta T_{RX,I}$ and $\Delta T_{RX,Q}$ are defined as the sum and difference of $\Delta T_{RX,v}$ and $\Delta T_{RX,h}$, respectively.

Using (3.4) gives $\Delta T_{RX,Q} = 0$. A more realistic description distinguishes between the calibration sources in the v and h channels, i.e.

$$\begin{aligned} \Delta T_{RX,v} &= \frac{T_{Hv} \hat{T}_{Cv} - T_{Cv} \hat{T}_{Hv}}{T_{Hv} - T_{Cv}}, \\ \Delta T_{RX,h} &= \frac{T_{Hh} \hat{T}_{Ch} - T_{Ch} \hat{T}_{Hh}}{T_{Hh} - T_{Ch}}, \end{aligned} \tag{3.5}$$

so that $\Delta T_{RX,Q}$ is nonzero. This distinction also complicates the expression for $\Delta T_{RX,U}$.

The radiometer calibration process is such that $\Delta T_{RX,I}$, $\Delta T_{RX,Q}$, and $\Delta T_{RX,U}$ are slowly varying (e.g., over a period of many minutes or more) compared with the radiometer integration time, τ. Even if estimates of the calibration parameters (e.g., $\hat{T}_{Cv}$) are obtained as often as several times per τ, it can be assumed that the predictable thermal environment of space permits extensive averaging of those estimates to yield better estimates. Therefore, for retrieving T_v, T_h, and T_Q measured in a single radiometer measurement cycle or even many cycles, we can consider $\Delta T_{RX,I}$, $\Delta T_{RX,Q}$, and $\Delta T_{RX,U}$ to be constants. It is also anticipated that this averaging (and other post-launch calibration activities) will reduce $\Delta T_{RX,I}$, $\Delta T_{RX,Q}$, and $\Delta T_{RX,U}$ to such low magnitude that they are negligible compared to the other error sources.

In the development of the above forward model, this chapter neglects the channel gains which are also estimated during data collection as part of the calibration process (see [2]). Although these gains and the uncertainties in them are relevant, they are omitted in this chapter, leaving their analysis for future work.

The quantities $\Delta T_{sys,I}$, $\Delta T_{sys,Q}$, and $\Delta T_{sys,U}$ in (3.1) through (3.3) are zero-mean, Gaussian random variables which correspond to the usual noise equivalent ΔT ($NE\Delta T$) of radiometric measurements [6]. They fluctuate significantly from one radiometer measurement cycle to the next. From their definition in (A.29) through (A.31), we see that they have the same covariance matrix as $\hat{T}_{sys,I}$, $\hat{T}_{sys,Q}$, and $\hat{T}_{sys,U}$ (as well as $\hat{T}_{Ia}$, $\hat{T}_{Qa}$, and $\hat{T}_{Ua}$), given in (3.6). In (3.6), $N \equiv 2B\tau$, where B is the sensor bandwidth (about 20 MHz for Aquarius and Hydros).

$$\begin{bmatrix} Var(\Delta T_{sys,I}) & Cov(\Delta T_{sys,I}, \Delta T_{sys,Q}) & Cov(\Delta T_{sys,I}, \Delta T_{sys,U}) \\ & Var(\Delta T_{sys,Q}) & Cov(\Delta T_{sys,Q}, \Delta T_{sys,U}) \\ & & Var(\Delta T_{sys,U}) \end{bmatrix} =$$

$$\frac{1}{N} \cdot \begin{bmatrix} T_{sys,I}^2 + T_{sys,Q}^2 + T_{sys,U}^2 & 2T_{sys,I}T_{sys,Q} & 2T_{sys,I}T_{sys,U} \\ & T_{sys,I}^2 + T_{sys,Q}^2 - T_{sys,U}^2 & 2T_{sys,Q}T_{sys,U} \\ & & T_{sys,I}^2 - T_{sys,Q}^2 + T_{sys,U}^2 \end{bmatrix} \quad (3.6)$$

For future reference, the means of the calibrated measurements are:

$$\begin{aligned} T_{Ia} \equiv <\hat{T}_{Ia}> &= \quad T_I \quad + \Delta T_{RX,I} \quad , \\ T_{Qa} \equiv <\hat{T}_{Qa}> &= +T_Q \cos 2\Omega + T_U \sin 2\Omega + \Delta T_{RX,Q} \quad , \\ T_{Ua} \equiv <\hat{T}_{Ua}> &= -T_Q \sin 2\Omega + T_U \cos 2\Omega + \Delta T_{RX,U} \quad . \end{aligned} \tag{3.7}$$

3.3 Rotation Correction Technique

In this section, PRC is reviewed. This is done in the context of the forward model summarized in Section 3.2.

3.3.1 Estimation of T_Q

Yueh's model [1] does not include any of the Δ terms in (3.1) through (3.3). By noting that T_U is much smaller than T_Q in natural earth scenes, he proposes to solve (3.2) and (3.3) for T_Q by also neglecting the terms with T_U. By assuming T_U and all the Δ quantities are zero, squaring both sides of (3.2) and (3.3), adding the two results, and then solving for T_Q, we obtain Yueh's proposed estimate,

$$\hat{T}_Q = \sqrt{\hat{T}_{Qa}^2 + \hat{T}_{Ua}^2}, \tag{3.8}$$

where we ignore the negative root since T_Q is positive in geophysical circumstances. In reality of course, T_U and all the Δ quantities are nonzero and constitute the error sources of the correction technique. Nevertheless, as demonstrated by the error analysis below, this equation provides a good estimate of T_Q.

3.3.2 Estimation of T_v and T_h

With $\hat{T}_Q$ from (3.8) and $\hat{T}_{Ia}$ from (3.1), we can also find $\hat{T}_v$ and $\hat{T}_h$ as $(\hat{T}_{Ia} \pm \hat{T}_Q)/2$. An error analysis of $\hat{T}_v$ and $\hat{T}_h$ is pursued in Section 3.5.

Yueh proposed an alternate method of estimating $\hat{T}_v$ and $\hat{T}_h$ (though it is straightforward to show its equivalence to estimating $\hat{T}_v$ and $\hat{T}_h$ via $(\hat{T}_{Ia} \pm \hat{T}_Q)/2$). First, to estimate Ω, we divide (3.3) by (3.2), then solve for Ω. With the error sources

(T_U and all the Δ quantities) assumed to be zero, this yields

$$\hat{\Omega} = \frac{1}{2}\tan^{-1}\frac{-\hat{T}_{Ua}}{\hat{T}_{Qa}} \tag{3.9}$$

as an estimate of the angle of polarization rotation. Then assuming the error sources are zero and solving (A.35) and (A.36) for T_v and T_h, respectively, yields

$$\hat{T}_v = \hat{T}_{va} + \hat{T}_Q \sin^2\hat{\Omega}, \tag{3.10}$$

$$\hat{T}_h = \hat{T}_{ha} - \hat{T}_Q \sin^2\hat{\Omega}, \tag{3.11}$$

which are the corrected forms of Eqs. (15) and (16) in [1], where $\hat{T}_Q$ is given in (3.8).

3.4 Analysis of $\hat{T}_Q$

The next task is to determine the pdf, mean, and variance of the estimate $\hat{T}_Q = \sqrt{\hat{T}_{Qa}^2 + \hat{T}_{Ua}^2}$. We use the mean and variance to calculate the RMSE. $\hat{T}_{Qa}$ and $\hat{T}_{Ua}$ are already well characterized. They are Gaussian (at least to a very good approximation) with means given in (3.7) and with variances and covariance given in (3.6).

3.4.1 Rotation of Variables

$\hat{T}_{Qa}$ and $\hat{T}_{Ua}$ are correlated, but we can rotate coordinates such that we have uncorrelated quantities. Define

$$\begin{bmatrix}\hat{Z}\\ \hat{W}\end{bmatrix} = \frac{1}{\sqrt{T_{sys,Q}^2 + T_{sys,U}^2}}\begin{bmatrix}T_{sys,Q} & T_{sys,U}\\ -T_{sys,U} & T_{sys,Q}\end{bmatrix}\begin{bmatrix}\hat{T}_{Qa}\\ \hat{T}_{Ua}\end{bmatrix}, \tag{3.12}$$

This is useful because $\sqrt{\hat{Z}^2 + \hat{W}^2} = \sqrt{\hat{T}_{Qa}^2 + \hat{T}_{Ua}^2} = \hat{T}_Q$. If we assume that $\hat{T}_{Qa}$ and $\hat{T}_{Ua}$ are *jointly* normal, then $\hat{Z}$ and $\hat{W}$ are also jointly normal and it is straightforward to show that $\hat{Z}$ and $\hat{W}$ are uncorrelated and have the following means and variances:

$$
\begin{aligned}
<\hat{Z}> &= \frac{T_{sys,Q}T_{Qa} + T_{sys,U}T_{Ua}}{\sqrt{T_{sys,Q}^2 + T_{sys,U}^2}} \equiv Z, & (3.13)\\
<\hat{W}> &= \frac{T_{sys,Q}T_{Ua} - T_{sys,U}T_{Qa}}{\sqrt{T_{sys,Q}^2 + T_{sys,U}^2}} \equiv W, & (3.14)\\
Var(\hat{Z}) &= \frac{T_{sys,I}^2 + T_{sys,Q}^2 + T_{sys,U}^2}{N} \equiv \sigma_Z^2, & (3.15)\\
Var(\hat{W}) &= \frac{T_{sys,I}^2 - T_{sys,Q}^2 - T_{sys,U}^2}{N} \equiv \sigma_W^2. & (3.16)
\end{aligned}
$$

The pdf of $\hat{T}_Q$ is given by using Z, W, σ_Z, and σ_W in (2) of [19], restated here in terms of the current problem:

$$
\begin{aligned}
f_{\hat{T}_Q}(\hat{T}_Q) &= \frac{\hat{T}_Q}{\sigma_Z \sigma_W} e^{-\frac{\hat{T}_Q^2+2Z^2}{4\sigma_Z^2}} e^{-\frac{\hat{T}_Q^2+2W^2}{4\sigma_W^2}} \cdot \\
&\quad \sum_{j=-\infty}^{\infty} I_j\left(a\hat{T}_Q^2\right) I_{2j}\left(d\hat{T}_Q\right) \cos 2j\psi & (3.17)
\end{aligned}
$$

for $\hat{T}_Q > 0$ and 0 otherwise, where

$$
a \equiv \frac{\sigma_Z^2 - \sigma_W^2}{4\sigma_Z^2\sigma_W^2}, \quad d^2 \equiv \frac{Z^2}{\sigma_Z^4} + \frac{W^2}{\sigma_W^4}, \quad \text{and} \quad \tan\psi \equiv \frac{W\sigma_Z^2}{Z\sigma_W^2},
$$

and the I_j are modified Bessel functions of the first kind and order j.

3.4.2 Simple Result by Assuming $\sigma_Z^2 \approx \sigma_W^2$

An attempt was made to find the first and second moments of $\sqrt{\hat{Z}^2 + \hat{W}^2}$ analytically, but failed, even though the pdf is known. Fortunately, a small approximation leads to simple and accurate formulas, as now shown.

$T_{sys,Q}^2 + T_{sys,U}^2$ has a worst case maximum value of about 4000 K^2 for Aquarius; in more extreme cases it might reach 10,000 K^2 (this is for L-band radiometers with incidence angles less than about 50°). But this is small compared to $T_{sys,I}^2$, which has a value of $660,000$ K^2 for typical Aquarius parameters ($T_{RX,I} = 620$ K and $T_I = 190$ K). Therefore, $\sigma_Z^2 \approx T_{sys,I}^2/N$ and $\sigma_W^2 \approx T_{sys,I}^2/N$.

If we define $\sigma^2 \equiv T_{sys,I}^2/N$ and use $\sigma_Z^2 \approx \sigma^2$ and $\sigma_W^2 \approx \sigma^2$, then $\hat{T}_Q$ is the root of the sum of the squares (RSS) of two independent Gaussians with the same variance and with nonzero, unequal means. Note that σ is the $NE\Delta T$ for the total signal (first Stokes parameter).

With this approximation, the pdf, mean, and variance of $\hat{T}_Q$ can be described as functions of just σ and m where $m^2 \equiv Z^2 + W^2 = T_{Qa}^2 + T_{Ua}^2$. In terms of the original parameters,

$$\begin{aligned} m^2 = \quad & T_Q^2 + T_U^2 + \Delta T_{RX,Q}^2 + \Delta T_{RX,U}^2 \\ & +2\cos 2\Omega(+T_Q \Delta T_{RX,Q} + T_U \Delta T_{RX,U}) \\ & +2\sin 2\Omega(-T_Q \Delta T_{RX,U} + T_U \Delta T_{RX,Q}). \end{aligned} \tag{3.18}$$

By either [20] or by (1) of [19], the density of $\hat{T}_Q$ is then

$$f_{\hat{T}_Q}(\hat{T}_Q) = \frac{\hat{T}_Q}{\sigma^2} e^{-\frac{\hat{T}_Q^2 + m^2}{2\sigma^2}} I_0\left(\frac{\hat{T}_Q m}{\sigma^2}\right), \hat{T}_Q > 0 \text{ (0 otherwise)}. \tag{3.19}$$

The mean and variance of $\hat{T}_Q$ are [21]:

$$< \hat{T}_Q >= \sigma\sqrt{\frac{\pi}{2}} e^{-\frac{m^2}{2\sigma^2}} {}_1F_1\left(\frac{3}{2}, 1; \frac{m^2}{2\sigma^2}\right), \tag{3.20}$$

$$Var(\hat{T}_Q) = 2\sigma^2 + m^2 - < \hat{T}_Q >^2, \tag{3.21}$$

where ${}_1F_1$ is the confluent hypergeometric function.

Eq. (3.20) corresponds with the first line of (3.10-12) in [22]; the second line shows that we can rewrite $< \hat{T}_Q >$ as

$$< \hat{T}_Q >= \sigma\sqrt{\frac{\pi}{2}} {}_1F_1\left(-\frac{1}{2}, 1; -\frac{m^2}{2\sigma^2}\right). \tag{3.22}$$

A difficulty with using either (3.20) or (3.22) is that for large τ, σ is small and the argument of ${}_1F_1$ has very large magnitude (e.g. 70,000 for the Aquarius $\theta = 28.7°$ case). Calculating the value of ${}_1F_1$ to high precision presents a huge computational burden when its argument is so large.

Fortunately, using (4B-9) in [22], (3.22) becomes

$$< \hat{T}_Q >= \sigma\sqrt{\frac{\pi}{2}} e^{-\frac{m^2}{4\sigma^2}} \left[(1 + \frac{m^2}{2\sigma^2}) I_0\left(\frac{m^2}{4\sigma^2}\right) + \frac{m^2}{2\sigma^2} I_1\left(\frac{m^2}{4\sigma^2}\right)\right], \tag{3.23}$$

which can be evaluated quickly.

The final simplification comes by examining plots of Monte Carlo results (see Section 3.4.3). These plots suggest that $Var(\hat{T}_Q) \approx \sigma^2$. Hypothesizing that $Var(\hat{T}_Q) \approx \sigma^2$ is correct and using this in (3.21) yields the simple formulas

$$\begin{aligned} < \hat{T}_Q > &\approx \sqrt{\sigma^2 + m^2}, && (3.24) \\ \Rightarrow Bias(\hat{T}_Q) &\approx \sqrt{\sigma^2 + m^2} - T_Q, && (3.25) \\ Var(\hat{T}_Q) &\approx \sigma^2 \text{ (our hypothesis)}, && (3.26) \\ \Rightarrow \text{RMSE of } \hat{T}_Q &= \sqrt{Var(\hat{T}_Q) + Bias^2(\hat{T}_Q)} && \\ &\approx \sqrt{2\sigma^2 + m^2 + T_Q^2 - 2T_Q\sqrt{\sigma^2 + m^2}}. && (3.27) \end{aligned}$$

These equations are the key result of this work.

Note that the pdf of $\hat{T}_Q$ given in (3.19) can be well approximated by a Gaussian pdf with the mean and variance of (3.24) and (3.26). Therefore, we are justified in ignoring higher moments hereafter and concerning ourselves with only the mean and variance (and the RMSE derived from them).

3.4.3 Validation of (3.24) through (3.27)

Numerical Equivalence of (3.23) and (3.24)

The final leap used to obtain (3.24) through (3.27) can be validated by showing that (3.24) matches (3.23). Maple™ finds the magnitude of the difference between (3.23) and (3.24) to be less than 20 nK in the Aquarius $\theta = 28.7°$ case, and less than 60 nK in the other Aquarius cases. (These are calculated with $T_U = \Delta T_{RX,Q} = 0.5$ K, $\Delta T_{RX,U} = 0$ K, and typical Aquarius values of $T_Q, T_I, T_{RX,I}$, and N for $-180° \leq \Omega \leq 180°$.)

Validation by Monte Carlo Simulation of Electric Field Model

The mean and variance of $\hat{T}_Q$ can also be found by Monte Carlo simulation. This can be done using (A.1) and (A.10) directly, thus avoiding all the approximations used in deriving (3.24) through (3.27) from (A.1) and (A.10).

The precise procedure is to generate N samples of a, b, E_v, and E_h, all independent of one another except $< E_v E_h >= T_U/2$. From these, N samples of

x and y are formed according to (A.1) and then squared and averaged to produce a single sample each of $\hat{T}_{sys,Q}$ and $\hat{T}_{sys,U}$ as in (A.10). To simulate the calibration process, $T_{RX,Q}$ is subtracted off while $\Delta T_{RX,Q}$ and $\Delta T_{RX,U}$ are added on, forming $\hat{T}_{Qa}$ and $\hat{T}_{Ua}$ as in (3.2) and (3.3). These are used in (3.8) to form a single sample of $\hat{T}_Q$. This entire procedure is repeated M times to form M independent samples of $\hat{T}_Q$. The empirical mean and variance of $\hat{T}_Q$ can then be calculated from these samples. This method gives no formulas but its results converge to the exact results as M increases.

The Monte Carlo results match analytic results from equations (3.25) through (3.27) very well, for many values of each of the parameters. Figures 3.1 through 3.4 show some of these results. The discrepancy can be attributed to the inherent imprecision in the Monte Carlo method.

3.5 Analysis of $\hat{T}_v$ and $\hat{T}_h$

This section analyzes $\hat{T}_v$ and $\hat{T}_h$, defined as $(\hat{T}_{Ia} \pm \hat{T}_Q)/2$. Using (3.1) and (3.24),

$$< \hat{T}_v > \;\approx\; \frac{1}{2}[T_I + \Delta T_{RX,I} + \sqrt{\sigma^2 + m^2}], \tag{3.28}$$

$$< \hat{T}_h > \;\approx\; \frac{1}{2}[T_I + \Delta T_{RX,I} - \sqrt{\sigma^2 + m^2}]. \tag{3.29}$$

3.5.1 Variance and RMSE of $\hat{T}_v$ and $\hat{T}_h$

Using (3.6) and (3.26)

$$\begin{aligned} Var(\hat{T}_v) &= \frac{1}{4}[Var(\hat{T}_{Ia}) + Var(\hat{T}_Q) + 2Cov(\hat{T}_{Ia}, \hat{T}_Q)] \\ &\approx \frac{1}{4}[\frac{2T_{sys,I}^2 + T_{sys,Q}^2 + T_{sys,U}^2}{N} + 2Cov(\hat{T}_{Ia}, \hat{T}_Q)]. \end{aligned} \tag{3.30}$$

Similarly,

$$Var(\hat{T}_h) \approx \frac{1}{4}[\frac{2T_{sys,I}^2 + T_{sys,Q}^2 + T_{sys,U}^2}{N} - 2Cov(\hat{T}_{Ia}, \hat{T}_Q)]. \tag{3.31}$$

Finding $Cov(\hat{T}_{Ia}, \hat{T}_Q)$ analytically appears to be intractable. But from (3.30) and (3.31) we see that $Cov(\hat{T}_{Ia}, \hat{T}_Q) = Var(\hat{T}_v) - Var(\hat{T}_h)$. We can therefore find $Cov(\hat{T}_{Ia}, \hat{T}_Q)$

numerically by subtracting the Monte Carlo estimates of $Var(\hat{T}_h)$ from the Monte Carlo estimates of $Var(\hat{T}_v)$. I studied such numerical results and found patterns, then hypothesized the following formula for $Cov(\hat{T}_{Ia}, \hat{T}_Q)$ from those patterns.

$$Cov(\hat{T}_{Ia}, \hat{T}_Q) = \frac{2T_{sys,I}}{N}\sqrt{T^2_{sys,Q} + T^2_{sys,U}}. \tag{3.32}$$

Using (3.32) in (3.30) and (3.31),

$$Var(\hat{T}_v) \approx \frac{2T^2_{sys,I} + 4T_{sys,I}\sqrt{T^2_{sys,Q} + T^2_{sys,U}} + T^2_{sys,Q} + T^2_{sys,U}}{4N}, \tag{3.33}$$

$$Var(\hat{T}_h) \approx \frac{2T^2_{sys,I} - 4T_{sys,I}\sqrt{T^2_{sys,Q} + T^2_{sys,U}} + T^2_{sys,Q} + T^2_{sys,U}}{4N}. \tag{3.34}$$

These, together with (3.28) and (3.29), give

$$\text{MSE of } \hat{T}_v \approx \frac{1}{4}[\sqrt{\sigma^2 + m^2} - T_Q + \Delta T_{RX,I}]^2 + \frac{2T^2_{sys,I} + 4T_{sys,I}\sqrt{T^2_{sys,Q} + T^2_{sys,U}} + T^2_{sys,Q} + T^2_{sys,U}}{4N}, \tag{3.35}$$

$$\text{MSE of } \hat{T}_h \approx \frac{1}{4}[\sqrt{\sigma^2 + m^2} - T_Q - \Delta T_{RX,I}]^2 + \frac{2T^2_{sys,I} - 4T_{sys,I}\sqrt{T^2_{sys,Q} + T^2_{sys,U}} + T^2_{sys,Q} + T^2_{sys,U}}{4N}. \tag{3.36}$$

The RMSE of estimated T_v and T_h are the positive square roots of these equations. They do not appear to simplify further, although the variances can be approximated as $\sigma^2/2$.

3.5.2 Plots

Figures 3.1 through 3.4 illustrate the bias (top), standard deviation (STD) (middle), and RMSE (bottom) for estimated T_Q, T_v, and T_h, as functions of Ω. The analytic results (given using (3.25) through (3.27), (3.28) through (3.29), and (3.33) through (3.36)) are plotted as solid lines. The Monte Carlo results are plotted as symbols.

Figures 3.1 and 3.2 were computed using T_I, T_Q [17], $T_{RX,I}$, and τ values typical of the innermost and outermost of the three Aquarius beams, respectively, while Fig. 3.3 uses values typical of the Hydros radiometer. The particular values of $T_U, T_{RX,Q}$, $\Delta T_{RX,I}$, and $\Delta T_{RX,Q}$ were chosen arbitrarily within their expected ranges. All values are given at the tops of the figures. The Monte Carlo results were generated as described above (Section 3.4.3).

The discrepancies between the analytic and the Monte Carlo results decrease as M increases. But it is difficult to increase M: generating a plot such as Fig. 3.1 currently requires days of CPU time. Another option is to generate the Monte Carlo samples using the Gaussian approximation (see Section A.3 in Appendix A). That is, rather than generating samples of the electric field, samples of $\hat{T}_{Ia}$, $\hat{T}_{Qa}$, and $\hat{T}_{Ua}$ themselves are generated as Gaussian random variables, with the means, variances, and covariances summarized in Section 3.2. This method, though not quite as exact, is many orders of magnitudes faster, allowing much larger M and more data points. Examples of the results obtained thereby are shown in Figs. 3.4-3.5.

3.6 Insights from the Equations

With the forward model developed in this chapter, there are five sources of error to be considered in polarization rotation correction: T_U, $\Delta T_{RX,I}$, $\Delta T_{RX,Q}$, $\Delta T_{RX,U}$, and $NE\Delta T$ (manifested as σ). In this section some effects of these error sources are studied, using the equations derived above. Note that Ω influences the expression of the error sources but is not an error source in itself.

3.6.1 The Insignificance of T_U

Examining (3.18), we see that the first term is desired while the rest are sources of bias. However, T_Q is more than an order of magnitude larger than the other components of (3.18); therefore we can neglect terms without T_Q, resulting in

$$m^2 \approx T_Q^2 + 2T_Q(\Delta T_{RX,Q}\cos 2\Omega - \Delta T_{RX,U}\sin 2\Omega). \tag{3.37}$$

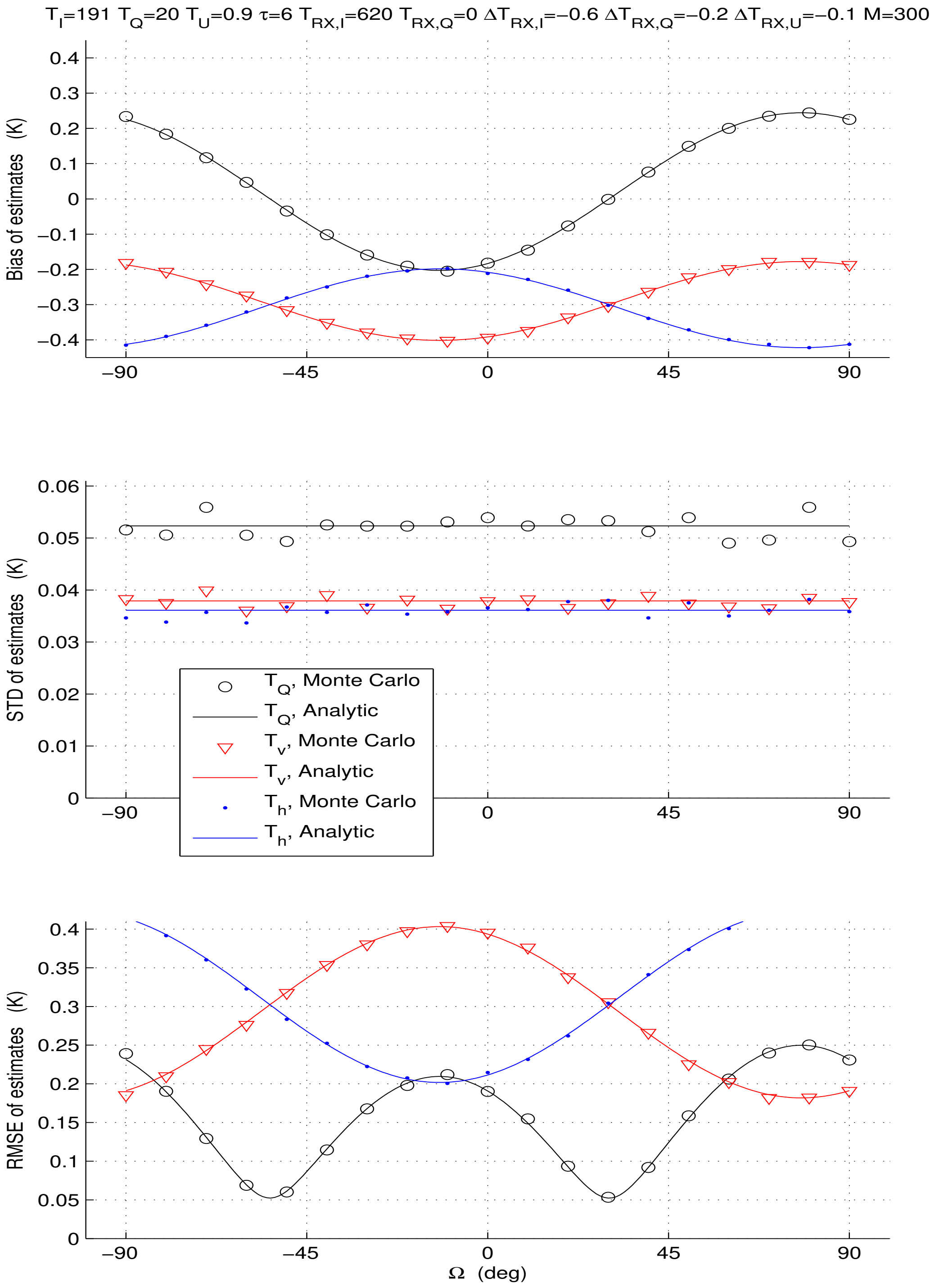

Figure 3.1: Bias (top), STD (center), and RMSE (bottom) of $\hat{T}_Q$, $\hat{T}_v$, and $\hat{T}_h$ as functions of Ω, with T_I, T_Q, τ, and $T_{RX,I}$ chosen to be typical of the Aquarius $\theta = 28.7°$ beam over ocean. The values of the remaining parameters ($T_{RX,Q}$, $\Delta T_{RX,I}$, $\Delta T_{RX,Q}$, and $\Delta T_{RX,U}$) were chosen arbitrarily within expected ranges.

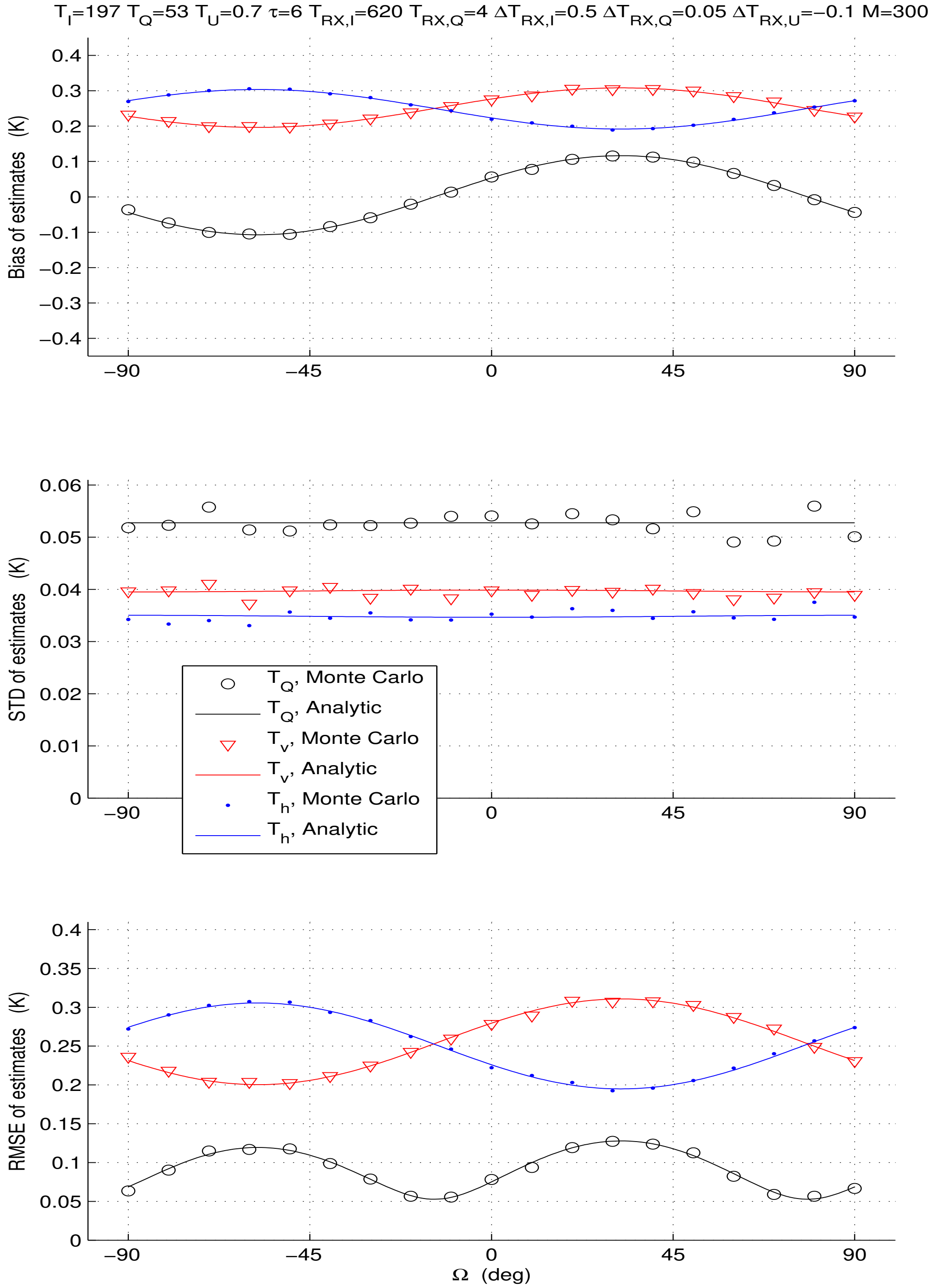

Figure 3.2: Same as Fig. 3.1 except for the Aquarius $\theta = 45.6^\circ$ beam. (Different choices of $T_{RX,Q}$, $\Delta T_{RX,I}$, $\Delta T_{RX,Q}$, and $\Delta T_{RX,U}$ were also used, to demonstrate the variety of possible behavior in the error.)

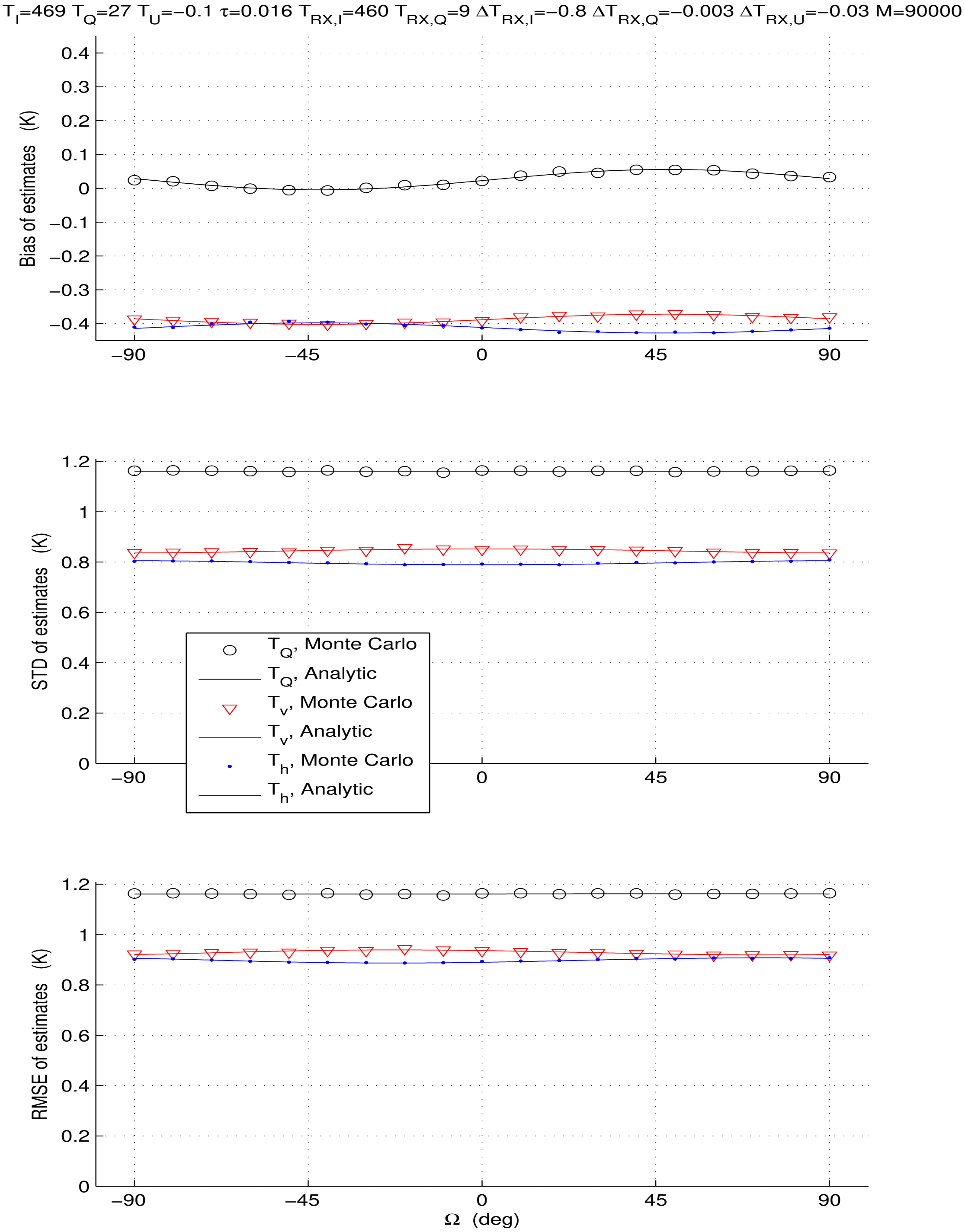

Figure 3.3: Same as Fig. 3.1 except for the Hydros soil moisture sensing mission. (Different choices of $T_{RX,Q}$, $\Delta T_{RX,I}$, $\Delta T_{RX,Q}$, and $\Delta T_{RX,U}$ were also used, to demonstrate the variety of possible behavior in the error.)

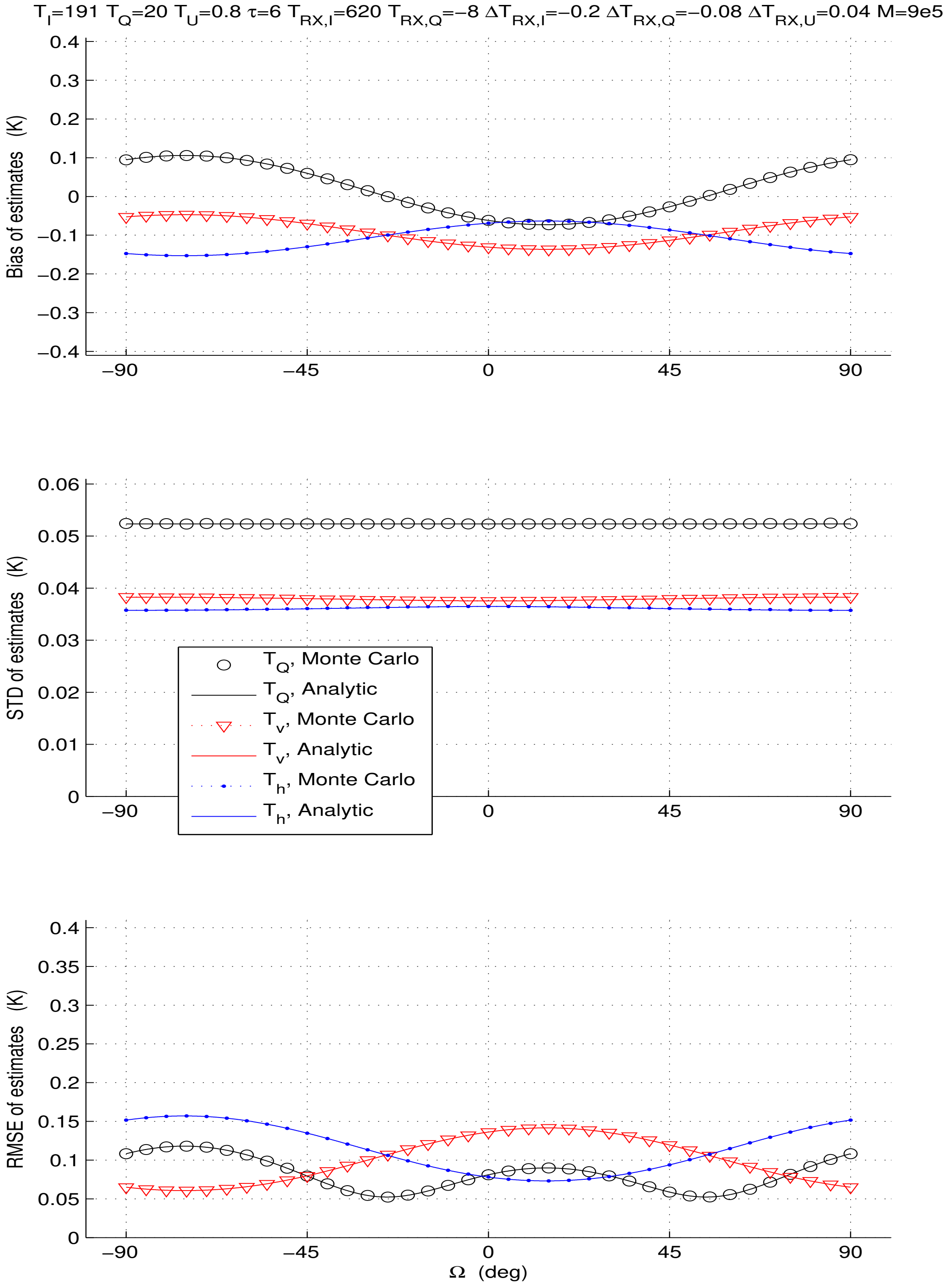

Figure 3.4: Same as Fig. 3.1 except with the Monte Carlo results generated using the Gaussian approximation, thus allowing much larger M. (Different choices of $T_{RX,Q}$, $\Delta T_{RX,I}$, $\Delta T_{RX,Q}$, and $\Delta T_{RX,U}$ were also used, to demonstrate the variety of possible behavior in the error.)

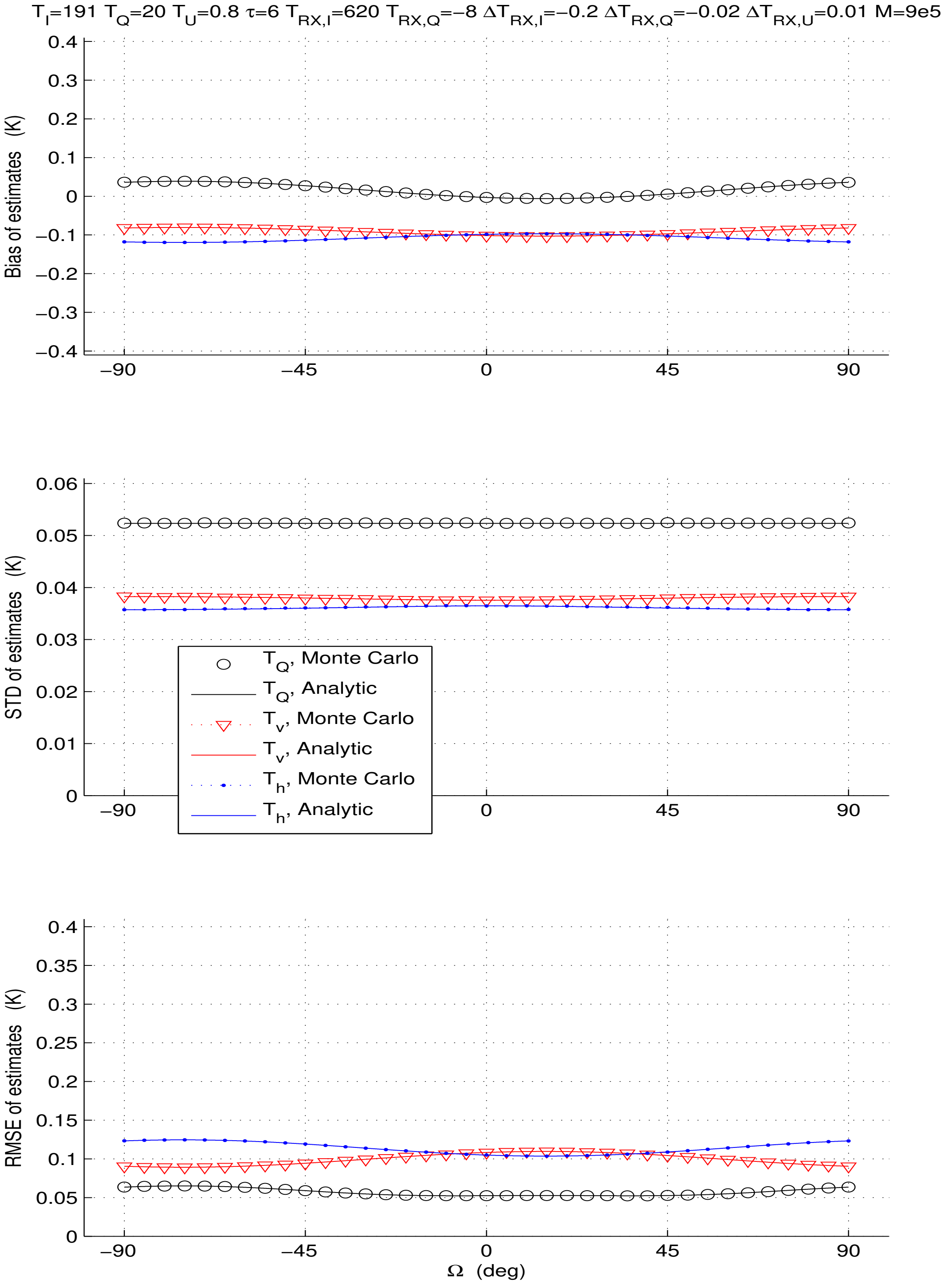

Figure 3.5: Same as Fig. 3.4 except that $\Delta T_{RX,I}$, $\Delta T_{RX,Q}$, $\Delta T_{RX,U}$ are made small compared to σ, which is the result anticipated from careful post-launch calibration. As a consequence, the dependence on Ω is weak in this figure.

This eliminates T_U from the equations, which suggests that natural T_U is not a significant error source in polarization rotation correction, at least at L-band.

We can numerically examine the significance of T_U as follows. We set the unknown error sources ($\Delta T_{RX,I}$, $\Delta T_{RX,Q}$, and $\Delta T_{RX,U}$) to zero but retain σ since it is known. Then we plot RMSE as a function of natural T_U. The results are shown in Fig. 3.6 for typical parameters of Aquarius beams. The RMSE at $T_U = 0$ is due to $NE\Delta T$ through σ. We can discern that the error caused by natural T_U is negligible compared to $NE\Delta T$ when $|T_U| < 1.5$ K.

Natural T_U is reported to have a maximum magnitude of about 1.5 K over the oceans at intermediate and high wind speeds, 10.7 GHz, and an incidence angle of 50° [23]. Extensive measurements at L-band have not been made, but one group reports amplitudes of less than 1 K over wind-driven ocean [24]. These measurements, combined with Fig. 3.6, suggest that natural T_U is not a significant error source for the Aquarius mission. This further suggests that the error allocation for "Other (Wind)" in the Aquarius error budget (on p. 8 of [14]) can be significantly reduced.

Plotted at the right of Fig. 3.6 are the results for typical parameters of a soil moisture sensing mission such as the canceled Hydros mission. The short integration time of the conical scanning radiometer results in $NE\Delta T$ being so large that the effects of natural T_U are negligible for $|T_U| < 5$ K.

3.6.2 The Optimal Ω Value

Examining (3.18) shows that near $\Omega = 0°$, the effect of $\Delta T_{RX,Q}$ is amplified compared to the effect of $\Delta T_{RX,U}$, because of T_Q being so much larger than T_U. Similarly, the effect of $\Delta T_{RX,U}$ is amplified near $\Omega = 45°$. Consequently, if $|\Delta T_{RX,Q}|$ is significantly larger than $|\Delta T_{RX,U}|$, then the RMSE of $\hat{T}_Q$ is minimum near $\Omega = 45°$. Likewise, if $|\Delta T_{RX,Q}|$ is significantly smaller than $|\Delta T_{RX,U}|$, the RMSE of $\hat{T}_Q$ is minimum near $\Omega = 0°$. See Figs. 3.1-3.4 for examples.

If $\Delta T_{RX,Q}$ and $\Delta T_{RX,U}$ have the same magnitude and sign, RMSE is minimum near $\Omega = 22.5°$. If they have the same magnitude but opposite sign, RMSE is minimum near $\Omega = -22.5°$. But in all these cases, if the magnitude of both is less

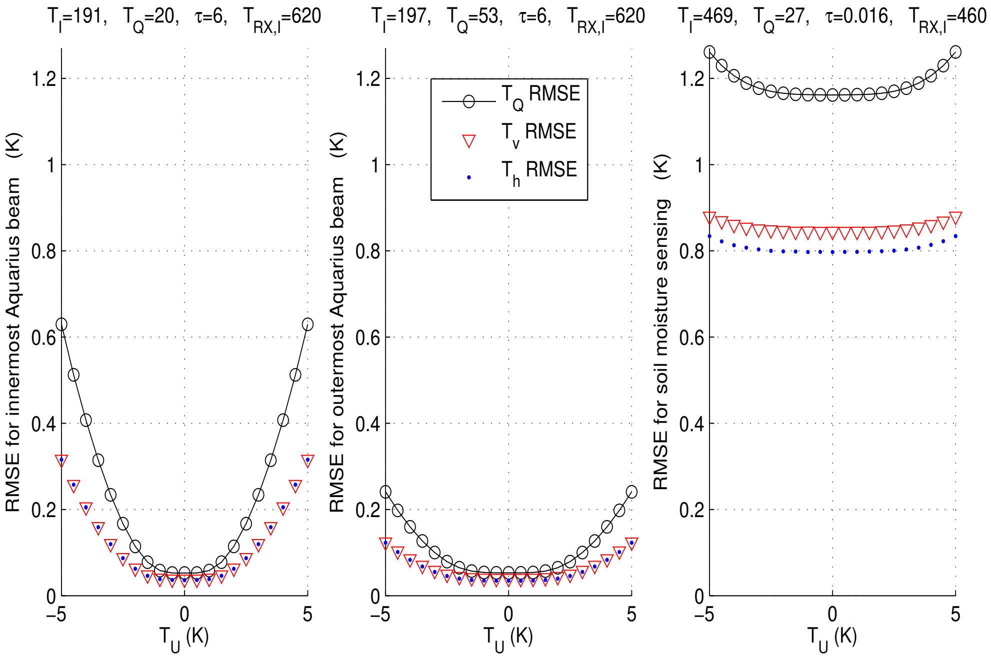

Figure 3.6: RMSE (from the analytically derived formulas) of $\hat{T}_Q$, $\hat{T}_v$, and $\hat{T}_h$ as functions of T_U, with $\Delta T_{RX,I}$, $\Delta T_{RX,Q}$, and $\Delta T_{RX,U}$ set to zero. T_I, T_Q, τ, and $T_{RX,I}$ are typical of the Aquarius $\theta = 28.7°$ beam (left), the Aquarius $\theta = 45.6°$ beam (center), and the Hydros soil moisture sensing mission (right).

than $\sigma/2$ then RMSE is approximately constant with respect to Ω (and is $\approx \sigma$ for $\hat{T}_Q$), as illustrated in Fig. 3.5.

Because $\Delta T_{RX,Q}$ and $\Delta T_{RX,U}$ are unknown (at least until instrument fabrication and initial calibration), there is no basis for claiming *a priori* that RMSE is better at any one value of Ω than at any other value. This should correct the notion that it is best to sense the land or ocean at dawn because of low free electron content in the atmosphere (and hence small Ω), an assumption used in the design of the Hydros mission. We note that there may be other good reasons for sensing at dawn, such as the better known temperature profile of the atmosphere and planetary surface.

3.6.3 Negligible Error Contribution of PRC

If $\Delta T_{RX,I}$, $\Delta T_{RX,Q}$, and $\Delta T_{RX,U}$ are reduced to insignificance through post-launch calibration, then the overall RMSE reduces to the $NE\Delta T$ that exists regardless of polarization rotation. To see this analytically, let $\Delta T_{RX,I} = \Delta T_{RX,Q} = \Delta T_{RX,U} = 0$ and also let $T_U = 0$ since we know its effect is not large. Then m^2 reduces to T_Q^2, and using a binomial expansion of (3.25),

$$Bias(\hat{T}_Q) \approx \sqrt{\sigma^2 + T_Q^2} - T_Q \;\approx\; \frac{\sigma^2}{2T_Q} \tag{3.38}$$

$$\Rightarrow \text{RMSE of } \hat{T}_Q \approx \sqrt{\sigma^2 + \frac{\sigma^4}{4T_Q^2}} \;\approx\; \sigma. \tag{3.39}$$

The validity of these approximations is easily confirmed by numerical examples using Aquarius and Hydros parameters.

Similar analysis shows that the RMSE of $\hat{T}_v$ and $\hat{T}_h$ reduces to $\sigma/2$. Hence, error allocation for ionospheric effects can be greatly reduced [14].

3.7 Conclusions

This chapter extends the forward model of polarization rotation to include the random nature of radiation, radiometer channel noises, and (to first order) calibration. In particular, it derives the means, variances, and covariances of the first three Stokes parameters, T_I, T_Q, and T_U, (or their modified counterparts, T_v, T_h, and T_U) as measured by radiometers.

There are several known limitations to this forward model. First, it ignores the uncertainties in channel gains which remain after the calibration process. Second, it ignores antenna sidelobe contributions to the apparent brightness temperature, which undergo a different amount of polarization rotation than the main beam contribution. Third, it ignores the mixing of the scene Stokes parameters by the antenna (i.e. the antenna cross-pol patterns). This chapter ignores these effects for tractability and because their effects are projected to be smaller than those effects which are included.

Using the forward model just described, this chapter analyzes the errors in using polarimetric measurements to correct for polarization rotation as proposed in [1]. Closed form equations have been derived for the bias, STD, and RMSE of estimated T_Q (see equations (3.25)-(3.27)) and similar expressions for T_v, and T_h. These equations match numerical results obtained by Monte Carlo simulation of our original electric field model. These equations are the key results of this work since they allow more accurate error analysis and error budgeting than has been possible previously.

This analysis indicates several things about the five sources of error. First, the natural third Stokes parameter (of the magnitude expected at L-band for most natural Earth scenes) is an insignificant source of error compared to $NE\Delta T$ (for $\tau \leq 6$ sec over ocean).

Second, the dependence of RMSE on rotation angle is determined by residual errors from the calibration process, $\Delta T_{RX,Q}$ and $\Delta T_{RX,U}$. Since these residuals are unknown (by definition), I cannot predict the dependence of RMSE on rotation angle (such as whether or not $\Omega = 0°$ is the optimum angle). But if post-launch calibration reduces $\Delta T_{RX,Q}$ and $\Delta T_{RX,U}$ to the level of $NE\Delta T$ or less, then the dependence of RMSE on Ω is weak.

Third, if $\Delta T_{RX,I}$, $\Delta T_{RX,Q}$, and $\Delta T_{RX,U}$ are reduced to insignificance through post-launch calibration, then overall RMSE reduces to the $NE\Delta T$ that exists regardless of polarization rotation. In other words, Yueh's method reduces the error incurred by polarization rotation to negligibility.

Chapter 4

Optimal Estimation of Calibration Parameters

In the previous chapter, probability theory was used to improve *analysis* of microwave radiometry. A probability theory-based approach can also improve the *usage* of microwave radiometry measurements, as demonstrated by a Bayesian approach in this chapter. Specifically, a probability theory-based approach is the means by which one can take full advantage of the redundancy in a set of measurements in order to reduce the inherent uncertainty (noise) in estimates of radiometer parameters. That is, probability theory enables optimal estimates.

The derivation of optimal estimates points the way to a fundamentally better way of reporting estimates and the uncertainties in them. The optimal solution to the estimation problem is a pdf for the set of radiometer parameters. The mean of the pdf provides the best (minimum average error) numerical estimates of the parameters. But the structure of the pdf provides much additional information: the uncertainties in the estimates (variances) and the *correlations* between uncertainties in pairs (covariances), triplets, etc. of the parameters. This information can be preserved and transmitted by reporting the full pdf for a set of parameters.

Rather than reporting a full pdf for a set of parameters, current practice in radiometry is often to report and use merely numerical estimates for parameters (such as the mean of the pdf). At best, variances are also reported. Reporting and utilizing a full pdf certainly requires more work. However, a purpose of this chapter is to demonstrate the advantages of a full pdf and to mention options for reporting and using it. This work is published as [5].

4.1 Introduction

A polarimetric radiometer measures the intensity of thermal emissions in at least three polarizations, represented by brightness temperature vector $[T_v \; T_h \; T_U]^T$. A simple forward model for the voltages recorded by one class of radiometer is

$$\begin{bmatrix} v_v \\ v_h \\ v_p \\ v_m \end{bmatrix} = \begin{bmatrix} G_{vv} & 0 & 0 \\ 0 & G_{hh} & 0 \\ G_{pv} & G_{ph} & G_{pU} \\ G_{mv} & G_{mh} & G_{mU} \end{bmatrix} \cdot \begin{bmatrix} T_v + T_1 \\ T_h + T_2 \\ T_U \end{bmatrix}, \tag{4.1}$$

where the G_{xx} are radiometer channel gains while T_1 and T_2 are the internal noises (represented by an equivalent noise temperatures) generated by the radiometer.

Radiometer calibration is the process of estimating the G_{xx}, T_1, and T_2 so that T_v, T_h, and T_U can be determined. Because the gains and noise temperatures can change rapidly during operation, this calibration is interleaved frequently between radiometer measurements of a scene. The calibration is accomplished by applying "known" inputs (hot and cold loads of temperature T_H and T_C) and measuring the voltage outputs of the radiometer.

Calibration of radiometers which measure the third Stokes parameter (T_U) requires an additional "known" input, T_{CN}, which is split and fed into both the vertical and horizontal channels so that the fluctuations in the electric fields of the two channels are correlated (simulating a third Stokes parameter input). See Fig. 1 in [2]. This technique was introduced in [2], for microwave radiometers which use a hybrid coupler to synthesize ±45° linear polarizations from vertical and horizontal signals. The noise-free forward model for the calibration measurements of such a radiometer is given below (Section 4.2). The algebraic method of [2] for estimating channel gains and temperatures using this noise-free forward model is summarized below in Section 4.3.

In the remainder of this work, we improve on the method of [2]. In Section 4.4 we expand the forward model to include the noise in the measurements and then solve the calibration problem using Bayes theorem rather than algebraically as was done in [2]. This results in a probability distribution function (pdf) for the calibration parameters.

This pdf itself is the best answer to the calibration problem and is explored in Section 4.6. However in Section 4.5, we compare numerical estimates extracted from this pdf with the algebraic estimates of [2]. We show that our estimates are optimal in the sense of minimizing root-mean-square-error (RMSE). They are unbiased, and their RMSE is approximately half the RMSE of the algebraic estimates.

4.2 Calibration Forward Problem

The noise free forward model for a single cycle of the full (Case 4) polarimetric radiometer calibration algorithm described in [2] can be written as

$$\begin{bmatrix} v_{v,C} & v_{v,H} & v_{v,CH} & v_{v,CN} \\ v_{h,C} & v_{h,H} & v_{h,CH} & v_{h,CN} \\ v_{p,C} & v_{p,H} & v_{p,CH} & v_{p,CN} \\ v_{m,C} & v_{m,H} & v_{m,CH} & v_{m,CN} \end{bmatrix} = \begin{bmatrix} G_{vv} & 0 & 0 \\ 0 & G_{hh} & 0 \\ G_{pv} & G_{ph} & G_{pU} \\ G_{mv} & G_{mh} & G_{mU} \end{bmatrix}$$

$$\times \begin{bmatrix} T_C + T_1 & T_H + T_1 & T_C + T_1 & T_C + \frac{1}{2}T_{CN} + T_1 \\ T_C + T_2 & T_H + T_2 & T_H + T_2 & T_C + \frac{1}{2}T_{CN} + T_2 \\ 0 & 0 & 0 & \pm T_{CN} \end{bmatrix}. \tag{4.2}$$

This forward model is (1) in [2], with o_v, o_h, o_p, and o_m given by (8) in [2], and with the vector $[T_v \quad T_h \quad T_U]^T$ replaced by the rightmost matrix in (4.2) in order to describe calibration measurements.

The four columns of the rightmost matrix in (4.2) represent the brightness temperature inputs used in the four types of calibration measurements (looks). The top row has temperature inputs to the vertically polarized (v) channel, the second row has inputs to the horizontally polarized (h) channel, and the third row has third Stokes parameter inputs (representing correlation between the vertically and horizontally polarized signals, T_U). .

We can describe the forward model as follows. On the left side of (4.2) are the 16 voltages measured in one radiometer calibration cycle: for each of the four calibration looks (cold load, hot load, mixed load, and cold load plus correlated noise load), the voltage outputs of the four polarimetric channels (v, h, p, and m) are measured. On the right side of (4.2) are ten calibration parameters which are unknown to some degree: eight radiometer gains plus two receiver noise temperatures T_1 and T_2. For this work we pretend that T_C, T_H, and T_{CN} are perfectly known. Optimal estimation of them from thermometer measurements, on-ground calibration, and the voltages on the left side of (4.2) is left for future work.

Various expansions of this model are possible. An additional column, the vector $[T_v + T_1 \qquad T_h + T_2 \qquad T_U]^T$, could be added to the right side, and a column of corresponding voltages to the left side. This vector contains the brightness of the scene under observation, whose estimation is the goal of radiometry. We have left off these columns in order to simplify the problem, focusing only on estimating the calibration parameters. However, our work can readily be extended to the larger problem. Other possible extensions include (a) allowing for nonzero G_{vh}, G_{vU}, G_{hv}, and G_{hU} in the gain matrix, rather than approximating them with zeros as done above (though this approximation is fairly accurate), (b) adding a small T_U term, generated by the radiometer, to the last row of the temperature matrix, (c) retrieving the radiometer hardware parameters that comprise the gain matrix entries rather than the entries themselves (this is accomplished in Section 4.7), and (d) adapting the method expounded in this work to other classes of radiometers, by using their forward models in place of (4.2).

4.3 Algebraic Estimation (the method of [2])

Estimating the ten calibration parameters on the right side of (4.2) by the method of [2] can be summarized as follows. It is algebraic, making no attempt to model the noise in the measurements.

G_{vv} is estimated using $v_{v,C}$ and $v_{v,H}$, which is the conventional cold/hot calibration method of non-polarimetric radiometers. From (4.2), the equations for these two voltages have the same form as equation (20) in [2], viz.,

$$\begin{bmatrix} v_{v,C} \\ v_{v,H} \end{bmatrix} = \begin{bmatrix} T_C & 1 \\ T_H & 1 \end{bmatrix} \begin{bmatrix} G_{vv} \\ G_{vv}T_1 \end{bmatrix}. \tag{4.3}$$

Solving for G_{vv}, the estimate is

$$\hat{G}_{vv} = \frac{v_{v,H} - v_{v,C}}{T_H - T_C} \tag{4.4}$$

as in (21) of [2]. T_1 is estimated by solving the same equations for T_1, yielding

$$\hat{T}_1 = \frac{-T_C v_{v,H} + T_H v_{v,C}}{v_{v,H} - v_{v,C}}. \tag{4.5}$$

Estimation of G_{hh} and T_2 from $v_{h,C}$ and $v_{h,H}$ is similar.

The gains G_{pv}, G_{ph}, and G_{pU} are found from (41) in [2], reproduced here (with o_p expanded and the minor correction that G_{pm} is replaced by G_{ph}):

$$\begin{bmatrix} v_{p,C} \\ v_{p,H} \\ v_{p,CH} \\ v_{p,CN} \end{bmatrix} = \begin{bmatrix} T_C & T_C & 0 & 1 \\ T_H & T_H & 0 & 1 \\ T_C & T_H & 0 & 1 \\ T_C + \frac{1}{2}T_{CN} & T_C + \frac{1}{2}T_{CN} & \pm T_{CN} & 1 \end{bmatrix} \begin{bmatrix} G_{pv} \\ G_{ph} \\ G_{pU} \\ G_{pv}T_1 + G_{ph}T_2 \end{bmatrix}. \tag{4.6}$$

Note that this also follows from (4.2). To obtain G_{pv}, G_{ph}, and G_{pU}, both sides of (4.6) are multiplied on the left by the inverse of the 4x4 temperature matrix. G_{mv}, G_{mh}, and G_{mU} are found similarly, using the measurements $v_{m,C}$, $v_{m,H}$, $v_{m,CH}$, and $v_{m,CN}$.

Note that the matrix in (4.6) has a condition number (for the 2-norm) of about 3000 (when we use load temperatures typical of NASA's near future Aquarius radiometer, $T_C = 288$ K and $T_H = T_{CN} = 800$ K [25]). This suggests that noise in the voltage measurements can disturb the estimates significantly. This may explain why our optimal estimates of the last six gains are the most improved compared to algebraic estimates (see Table 4.2). Also note that four of the available measurements, $v_{v,CH}$, $v_{v,CN}$, $v_{h,CH}$ and $v_{h,CN}$, are not utilized in this method. Nevertheless, the method of [2] works well.

4.4 Posterior Pdf, $p(\mathbf{m}|\mathbf{v})$

The algebraic estimation method sketched in the previous section is not unique. For example, G_{vv} could also be estimated using

$$\hat{G}_{vv} = \frac{v_{v,H} - v_{v,CH}}{T_H - T_C}. \tag{4.7}$$

A better estimate would be the average of (4.4) and (4.7), since the noise in each estimate is somewhat different and is therefore reduced by the averaging. Information on G_{vv} is also contained in the measurement $v_{v,CN}$, and similarly with the other nine calibration parameters. How can we combine all of the information in the voltages to get the best estimates of the ten calibration parameters? The answer is to approach the problem using probability theory, rather than algebraically. Using Bayes theorem, a pdf can express all the information on a parameter. This information can come from various measurements, from the noise-free forward model, and from a probabilistic description of the noise [9]. Let $\mathbf{v}$ be the vector[1] of measurements (voltages) in (4.2) and $\mathbf{m}$ be the vector of calibration parameters (the eight gains plus T_1 and T_2). Then Bayes' theorem tells us that the pdf for the parameters, given the voltages, is

$$p(\mathbf{m}|\mathbf{v}) = \frac{p(\mathbf{v}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{v})}. \tag{4.8}$$

The pdf $p(\mathbf{m})$ represents prior or external information on the model parameters. Several options are available for choosing an informative $p(\mathbf{m})$. A uniform pdf could be chosen, using known physical limits on each parameter. Measurements of these parameters by means other than the calibration voltages could also provide $p(\mathbf{m})$. A third option is to use historical values of the parameters from measurements of similar hardware. However, it is apparent in the results shown below that the calibration voltages provide high quality information on the parameters even without the use of an informative $p(\mathbf{m})$. Therefore, for simplicity, we choose a non-informative $p(\mathbf{m})$ – namely, a "flat" pdf represented by an arbitrary constant[2]. The denominator

---

[1] Vectors are denoted by boldface font in this chapter and are column vectors.

[2] The use of an arbitrary constant for a non-informative pdf can be considered the limit of a uniform pdf as its support is extended without bound in all directions.

of (4.8) is also a constant[3], so that

$$p(\mathbf{m}|\mathbf{v}) = cp(\mathbf{v}|\mathbf{m}), \tag{4.9}$$

where c is a constant.

4.4.1 Pdf for the Voltages Given Parameters, $p(\mathbf{v}|\mathbf{m})$

We now elaborate the probability distribution for voltages given a set of calibration parameters, $p(\mathbf{v}|\mathbf{m})$.

Noise Model

Equation (4.2) is a noise-free forward model for the 16 voltages from the calibration parameters. However, the temperature inputs in the last matrix of (4.2) are only mean values. Actual thermal emissions fluctuate. These random fluctuations can be treated as Gaussian noise (commonly called $NE\Delta T$) that is added to each "true" (mean value) temperature in the forward model. Therefore a forward model with noise included is

$$\begin{bmatrix} v_{v,C} & v_{v,H} & v_{v,CH} & v_{v,CN} \\ v_{h,C} & v_{h,H} & v_{h,CH} & v_{h,CN} \\ v_{p,C} & v_{p,H} & v_{p,CH} & v_{p,CN} \\ v_{m,C} & v_{m,H} & v_{m,CH} & v_{m,CN} \end{bmatrix} = \begin{bmatrix} G_{vv} & 0 & 0 \\ 0 & G_{hh} & 0 \\ G_{pv} & G_{ph} & G_{pU} \\ G_{mv} & G_{mh} & G_{mU} \end{bmatrix} \times$$
$$\begin{bmatrix} T_C + T_1 + n_1 & T_H + T_1 + n_3 & T_C + T_1 + n_5 & T_C + T_1 + \frac{1}{2}T_{CN} + n_7 \\ T_C + T_2 + n_2 & T_H + T_2 + n_4 & T_H + T_2 + n_6 & T_C + T_2 + \frac{1}{2}T_{CN} + n_8 \\ 0 & 0 & 0 & T_{CN} + n_9 \end{bmatrix}. \tag{4.10}$$

The noise term n_1 represents the total fluctuation away from $T_C + T_1$ during the first calibration sub-cycle. It is zero-mean Gaussian noise with variance equal to the mean temperature squared and then divided by the time bandwidth product, $\frac{(T_C+T_1)^2}{B\tau_c}$ (see Appendix B). n_2 through n_9 are similar[4].

---

[3] For fixed v, the denominator of (4.8) is obtained by integrating the numerator of (4.8), thus normalizing the right side of (4.8) so that $\int p(\mathbf{m}|\mathbf{v}) = 1$.

[4] Caveat: in a more detailed model, the zeros on the last row would be small noise terms, n_{10}, n_{11}, and n_{12}, representing the random amount of instantaneous correlation between the electric field

The noise terms n_1 through n_6 are all independent of one another (and of n_7 through n_9) for one or both of the following reasons: (a) they originate from different sources (the v channel hot and cold inputs, as well as the amplifiers producing T_1, are separate from the h channel inputs and the amplifiers producing T_2) and (b) they are realized during different calibration sub-cycles (i.e., are different realizations of the noise, and the rapidity of the fluctuations means that the realization during one interval is independent of the realization during the next interval). The noise terms n_7 through n_9 are not independent of one another since they are all originate (at least in part) from the correlated calibration source; they are treated in Appendix B.4.

Even though n_1 through n_6 are independent of one another, there is correlation among the voltages. For example, $v_{p,C}$ and $v_{m,C}$ are both functions of n_1 and n_2 and are therefore correlated. All the correlations that exist among the voltages can be summarized in a covariance matrix C. This C is derived in Appendix B. As a result, the probability distribution for $\mathbf{v}$ given $\mathbf{m}$ is a 16-dimensional Gaussian pdf, with mean given by the right hand side of (4.2) and denoted $\mathbf{g(m)}$:

$$p(\mathbf{v}|\mathbf{m}) = \frac{1}{\sqrt{(2\pi)^{16}\ |C|}} e^{-\frac{1}{2}(\mathbf{v}-\mathbf{g(m)})^T\ C^{-1}\ (\mathbf{v}-\mathbf{g(m)})}. \tag{4.11}$$

Eigendecomposition of Singular C

The covariance matrix C is a function of the calibration parameters, $\mathbf{m}$. Numerical calculations using arbitrary values for $\mathbf{m}$ show that although C is 16x16, its rank r is consistently only nine. From a theoretical standpoint, this corresponds to the nine noise sources on the right hand side of (4.10).

Because C is not full rank, it cannot be inverted. It also has $16 - r = 7$ eigenvalues with value zero so that its determinant $|C|$ is zero. How do we evaluate (4.11) in this situation?

---

in the v channel and the electric field in the h channel. But for simplicity, these small terms are treated as zero in this work.

Consider the eigendecomposition of C:[5]

$$C = [V \Lambda V^T] = \begin{bmatrix} V_1 \ V_2 \end{bmatrix} \ diag([\mathbf{\Lambda_1} \ \mathbf{0}]) \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}, \tag{4.12}$$

where the $r = 9$ columns of V_1 are eigenvectors with nonzero eigenvalues $\mathbf{\Lambda_1}$ while the seven columns of V_2 are eigenvectors with eigenvalues equal to zero; and $diag(\cdot)$ is a diagonal matrix with the elements of the argument along its diagonal. Matrix theory tells us that because C is symmetric, $V \equiv [V_1 V_2]$ is unitary [26]. Therefore, $V^T \mathbf{v} = \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \mathbf{v}$ is a rotation of the voltages. Consider a partition of these rotated voltages into $\mathbf{y} \equiv V_1^T \mathbf{v}$ and $\mathbf{x} \equiv V_2^T \mathbf{v}$. The covariance of $\mathbf{x}$ is

$$\begin{aligned} Cov(\mathbf{x}) &= \left\langle (\mathbf{x} - <\mathbf{x}>)(\mathbf{x} - <\mathbf{x}>)^T \right\rangle & (4.13) \\ &= \left\langle (V_2^T \mathbf{v} - <V_2^T \mathbf{v}>)(V_2^T \mathbf{v} - <V_2^T \mathbf{v}>)^T \right\rangle & (4.14) \\ &= V_2^T \left\langle (\mathbf{v} - <\mathbf{v}>)(\mathbf{v} - <\mathbf{v}>)^T \right\rangle V_2 & (4.15) \\ &= V_2^T Cov(\mathbf{v}) V_2 = V_2^T C V_2 = [0]_{7x7}, & (4.16) \end{aligned}$$

where the last step follows from the fact that the columns of V_2 are eigenvectors of C with eigenvalues equal to zero.

Because the variances of each of the x_i are zero, $\mathbf{x}$ is a constant vector. That is, for fixed $\mathbf{m}$, $\mathbf{x}$ will be the same for any realization of the voltages.

In this subsection we consider $\mathbf{m}$ to be given. Therefore we can find C and then V_2. We can also find the particular $\mathbf{v} = \mathbf{g}(\mathbf{m})$ corresponding to the voltages obtained when all nine noise sources happen to be zero. Therefore we can find $\mathbf{x}$ as $V_2^T \mathbf{g}(\mathbf{m})$. Since $\mathbf{x}$ is constant for any realization of the voltages, then all realizations of the voltages satisfy $V_2^T \mathbf{v} = V_2^T \mathbf{g}(\mathbf{m})$.

This last expression is a constraint when evaluating $p(\mathbf{v}|\mathbf{m})$. It says that the pdf is concentrated on a manifold (in the 16-dimensional space of voltages) defined by $V_2^T \mathbf{v} = V_2^T \mathbf{g}(\mathbf{m})$. For any $\mathbf{v}$ which does not satisfy this constraint (i.e., does not lie on the manifold), $p(\mathbf{v}|\mathbf{m}) = 0$.

---

[5] This is also the singular value decomposition (SVD) of C. But since C is symmetric, the SVD degenerates to an eigendecomposition, which is more convenient both numerically and analytically.

Next consider the rotated voltages $\mathbf{y} \equiv V_1^T\mathbf{v}$. By the same steps as in (4.16), we find that

$$\begin{aligned} Cov(\mathbf{y}) &= V_1^T C V_1 = V_1^T \begin{bmatrix} V_1 \; V_2 \end{bmatrix} \; diag([\mathbf{\Lambda_1} \; \mathbf{0}]) \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} V_1 && (4.17) \\ &= V_1^T V_1 diag(\mathbf{\Lambda_1}) V_1^T V_1 = diag(\mathbf{\Lambda_1}), && (4.18) \end{aligned}$$

which is nonzero. This new covariance matrix is invertible and its determinant is the product of the elements of $\mathbf{\Lambda_1}$. Therefore we can reduce the dimensionality of the Gaussian pdf $p(\mathbf{v}|\mathbf{m})$ from 16 to $r = 9$ by using $\mathbf{y}$ and its invertible covariance matrix in place of $\mathbf{v}$ and C. The mean of $\mathbf{y}$ is $V_1^T\mathbf{g}(\mathbf{m})$.

In summary, we have rotated and partitioned our measurements $\mathbf{v}$ into those which provide constraints ($\mathbf{x}$) and those which provide a smaller dimensional pdf ($\mathbf{y}$), so that (4.11) can be rewritten[6] as

$$p(\mathbf{v}|\mathbf{m}) = \begin{cases} \frac{1}{\sqrt{(2\pi)^9 \left|\prod_{i=1}^{9} \Lambda_i\right|}} e^{-\frac{1}{2}(\mathbf{v}-\mathbf{g}(\mathbf{m}))^T V_1 \cdot (diag(\mathbf{\Lambda_1}))^{-1} \cdot V_1^T (\mathbf{v}-\mathbf{g}(\mathbf{m}))} & \text{if } V_2^T\mathbf{v} = V_2^T\mathbf{g}(\mathbf{m}) \\ 0 & \text{if } V_2^T\mathbf{v} \neq V_2^T\mathbf{g}(\mathbf{m}). \end{cases}$$

In Appendix C, we derive V_2 analytically and show that the constraint $V_2^T\mathbf{v} = V_2^T\mathbf{g}(\mathbf{m})$ reduces to

$$G_{pv} = G_{vv}\frac{(v_{p,C}v_{h,H} - v_{h,C}v_{p,H})}{(v_{v,C}v_{h,H} - v_{h,C}v_{v,H})}, \quad G_{ph} = G_{hh}\frac{(v_{p,C}v_{v,H} - v_{v,C}v_{p,H})}{(v_{h,C}v_{v,H} - v_{v,C}v_{h,H})}, \qquad (4.21)$$

$$G_{mv} = G_{vv}\frac{(v_{m,C}v_{h,H} - v_{h,C}v_{m,H})}{(v_{v,C}v_{h,H} - v_{h,C}v_{v,H})}, \; G_{mh} = G_{hh}\frac{(v_{m,C}v_{v,H} - v_{v,C}v_{m,H})}{(v_{h,C}v_{v,H} - v_{v,C}v_{h,H})}, \qquad (4.22)$$

$$G_{mU} = G_{pU} \cdot \frac{G_{mv}G_{hh}v_{v,CN} + G_{mh}G_{vv}v_{h,CN} - G_{vv}G_{hh}v_{m,CN}}{G_{pv}G_{hh}v_{v,CN} + G_{ph}G_{vv}v_{h,CN} - G_{vv}G_{hh}v_{p,CN}}. \qquad (4.23)$$

---

[6] The exponent can also be written

$$\begin{aligned} -\frac{1}{2}\left(V_1^T\mathbf{v} - V_1^T\mathbf{g}(\mathbf{m})\right)^T \begin{bmatrix} \frac{1}{\Lambda_1} & & \\ & \ldots & \\ & & \frac{1}{\Lambda_9} \end{bmatrix} \left(V_1^T\mathbf{v} - V_1^T\mathbf{g}(\mathbf{m})\right) && (4.19) \\ = -\frac{1}{2}\sum_{i=1}^{9} \frac{\left([V(:,i)]^T \cdot (\mathbf{v} - \mathbf{g}(\mathbf{m}))\right)^2}{\Lambda_i}. && (4.20) \end{aligned}$$

4.5 Maximum A Posteriori / Maximum Likelihood Estimation

4.5.1 Theory

In Section 4.4.1 we demonstrated how to calculate $p(\mathbf{v}|\mathbf{m})$, see (4.21). From this we can find a distribution for the parameters of interest, $p(\mathbf{m}|\mathbf{v})$, by simply reversing the roles of input and output in the function and multiplying by a normalizing constant, as shown in (4.9)[7]. (We will not attempt to find the constant c since it is not necessary for finding estimates and since the pdf can be displayed in unnormalized form.) Explicitly,

$$p(\mathbf{m}|\mathbf{v}) = \frac{c}{\sqrt{(2\pi)^9 |\prod_{i=1}^{9} \Lambda_i|}} e^{-\frac{1}{2}(\mathbf{v}-\mathbf{g}(\mathbf{m}))^T V_1 \cdot (diag(\mathbf{\Lambda_1}))^{-1} \cdot V_1^T (\mathbf{v}-\mathbf{g}(\mathbf{m}))} \tag{4.24}$$

if (4.21)–(4.23) are satisfied, while

$$p(\mathbf{m}|\mathbf{v}) = 0$$

otherwise.

In general, the best answer to an estimation problem is a joint pdf on the variables of interest [3], in our case equation (4.24). A joint pdf usually contains much more information than merely reporting numbers and standard deviations for the variables. This is illustrated in Section 4.6.2. The increase in computing power of the last few decades enables us to begin to use pdfs (such as $p(\mathbf{m}|\mathbf{v})$) as inputs and outputs to algorithms, rather than simple estimates and their uncertainties. It is our hope that science and engineering will move in that direction. In this section, however, we follow tradition by reporting simple numerical estimates and their uncertainties (RMSE).

From the joint pdf, $p(\mathbf{m}|\mathbf{v})$, which numbers should we extract and report as estimates of the calibration parameters? Of all possible estimates of $\mathbf{m}$ from $\mathbf{v}$, the minimum mean-squared error (MMSE) estimate is the expected value (over $\mathbf{m}$) of $p(\mathbf{m}|\mathbf{v})$ [26]. Section 4.6.2 explains one way to calculate this expectation numerically.

---

[7] Due to the simplifying use of a constant for $p(\mathbf{m})$, we recognize that $p(\mathbf{m}|\mathbf{v})$ is identical to the likelihood function for $\mathbf{m}$, given $\mathbf{v}$.

However, a shortcut is available. Marginal 1-D and 2-D pdfs for any parameter or pair of parameters can be obtained by integrating $p(\mathbf{m}|\mathbf{v})$ with respect to the remaining parameters. Typical examples of these pdfs are shown in Figs. 4.1 through 4.4 (the generation of such figures is explained in Section 4.6). From such examples it appears that $p(\mathbf{m}|\mathbf{v})$ is unimodal and symmetric in most (if not all) cases. This is corroborated by the Gaussian structure of (4.24). These properties (unimodal and symmetric) signify that the mean of $p(\mathbf{m}|\mathbf{v})$ is the same as its mode (at least as a good approximation, if not exactly). Finding the mode turns out to be simpler than finding the mean; therefore we use the mode as our primary estimate. We examine the properties of the mode in detail in this section and return to consider the mean in Section 4.6.2.

$p(\mathbf{m}|\mathbf{v})$ is often called the posterior distribution, since it is the distribution for $\mathbf{m}$ *after* measuring data $\mathbf{v}$. The mode of $p(\mathbf{m}|\mathbf{v})$ is the set of parameters, $\mathbf{m}$, which maximizes the posterior distribution. Therefore it is referred to as the maximum a posteriori (MAP) estimate. In the current problem, since $p(\mathbf{m}|\mathbf{v})$ is equal to a constant times $p(\mathbf{v}|\mathbf{m})$, the MAP estimate is equivalent to what is called the maximum likelihood estimate, obtained by considering $p(\mathbf{v}|\mathbf{m})$ to be a function of $\mathbf{m}$ and finding the $\mathbf{m}$ which maximizes it.

Finding the $\mathbf{m}$ which maximizes $p(\mathbf{m}|\mathbf{v})$ is a standard multidimensional optimization problem. It is readily accomplished by a blackbox minimization algorithm such as MATLAB's fminsearch (though more advanced techniques could find it with less computation). The search can be initialized using the algebraic estimate of $\mathbf{m}$ (see Section 4.3). (Also, maximizing the log of $p(\mathbf{m}|\mathbf{v})$, rather than $p(\mathbf{m}|\mathbf{v})$ itself, causes the search to converge more quickly, about 3x speedup in the overall algorithm.) To limit the search to $\mathbf{m}$ which satisfy the constraint, we only search over the five parameters G_{vv}, G_{hh}, G_{pU}, T_1 and T_2. When the other five are needed, they are generated from the constraint equations, (4.21) through (4.23).

Table 4.1: Gain-related parameters typical of an L-Band radiometer and used in simulations

Radiometer hardware parameters	
detector sensitivities, c_v, c_h, c_p, and c_m	450 V/W
channel amplifier gain, G_1	1.8e7 W/W
gain imbalance, g	1.585
scattering parameter, s	0.7
α_e	0.934
bandwidth, B	20 MHz
Resulting gains	
G_{vv}	2.24e-6 V/K
G_{hh}	3.55e-6 V/K
G_{pv}	1.10e-6 V/K
G_{ph}	1.81e-6 V/K
G_{pU}	1.31e-6 V/K
G_{mv}	1.14e-6 V/K
G_{mh}	1.74e-6 V/K
G_{mU}	-1.31e-6 V/K

Note that the constraint equations dictate that estimates of G_{pv} and G_{mv} are 100% correlated with estimates of G_{vv}, estimates of G_{ph} and G_{mh} are 100% correlated with estimates of G_{hh}, and estimates of G_{pU} and of G_{mU} are also 100% correlated. In other words, a MAP estimate of any gain in the gain matrix is 100% correlated with the MAP estimates of the other members of the same column (see Fig. 4.4 for an illustration).

4.5.2 Simulation and Results

To compare MAP estimates with algebraic estimates, we simulate the estimation process as follows. Typical values of radiometer hardware parameters defined in [2], obtained from [25] and Table I of [2], are shown in Table 4.1. These are used in (17) of [2] to calculate typical values for the eight gains, G_{xx}, also shown in Table 4.1.

For typical T_1 and T_2 we use 310 K; these plus the eight G_{xx} in Table 4.1 are considered the "true" values in our simulations, $\mathbf{m}_{true}$. For T_C, T_H, and T_{CN} we use 288, 800, and 800 K, respectively. (These T_1, T_2, T_C, T_H, and T_{CN} are nominal values for NASA's Aquarius radiometer [25].)

From the above parameters we next generate simulated voltages. As discussed in Section 4.4.1, we can find $\mathbf{x}$ as $V_2^T\mathbf{g}(\mathbf{m})$. We also need realizations of $\mathbf{y}$. Since $\mathbf{y}$ is a linear combination of $\mathbf{v}$, it is Gaussian as is $\mathbf{v}$. We know that $\mathbf{y}$ has a mean of $V_1^T\mathbf{g}(\mathbf{m})$ and a covariance matrix of $diag(\mathbf{\Lambda_1})$ (see Section 4.4.1), so we can readily generate random realizations of it by adding the mean to zero-mean Gaussian noise with that covariance. Then with $\mathbf{x}$ and $\mathbf{y}$ in hand, we can form $\mathbf{v}$ by de-rotating (left multiplying by V) since

$$V\begin{bmatrix}\mathbf{y}\\ \mathbf{x}\end{bmatrix} = V\begin{bmatrix}V_1^T\mathbf{v}\\ V_2^T\mathbf{v}\end{bmatrix} = V\left(V^T\mathbf{v}\right) = \left(VV^T\right)\mathbf{v} = \mathbf{v}. \tag{4.25}$$

This can also be written to separate the voltages into their means and zero-mean noise:

$$\begin{aligned}\mathbf{v} &= V\begin{bmatrix}\mathbf{y}\\ \mathbf{x}\end{bmatrix} = [V_1 \quad V_2]\begin{bmatrix}V_1^T\mathbf{g}(\mathbf{m}) + N\left(0, diag(\mathbf{\Lambda_1})\right)\\ V_2^T\mathbf{g}(\mathbf{m})\end{bmatrix} && (4.26)\\ &= V_1[V_1^T\mathbf{g}(\mathbf{m}) + N\left(0, diag(\mathbf{\Lambda_1})\right)] + V_2V_2^T\mathbf{g}(\mathbf{m}) && (4.27)\\ &= \left(V_1V_1^T + V_2V_2^T\right)\mathbf{g}(\mathbf{m}) + V_1N\left(0, diag(\mathbf{\Lambda_1})\right), && (4.28)\end{aligned}$$

where $N\left(0, diag(\mathbf{\Lambda_1})\right)$ is a vector of zero-mean Gaussian random variables with covariance matrix $diag(\mathbf{\Lambda_1})$. The voltages generated in this manner satisfy the constraint: numerical tests show that $||V_2^T\mathbf{v} - V_2^T\mathbf{g}(\mathbf{m}_{true})|| \leq 1e-27$.

With the simulated voltages, we then use the method of [2], as summarized in Section 4.3, to algebraically estimate the ten calibration parameters. We repeat this for 10^6 different realizations of the voltages (means stay the same, noise changes). The bias of these 10^6 estimates is computed as the difference between their mean and $\mathbf{m}_{true}$. The standard deviation (STD) of these 10^6 estimates is also computed. Finally, the RMSE of these 10^6 estimates is computed as the root-sum-square of the bias and

Table 4.2: RMSE results for 10^6 estimates from voltages generated by typical $\mathbf{m}_{true}$. RMSE is given as a percent of the true parameter value.

Parameter	G_{vv}	G_{hh}	G_{pv}	G_{ph}	G_{pU}	G_{mv}	G_{mh}	G_{mU}	T_1	T_2
RMSE of algebraic estimates (%)	0.58	0.58	1.33	0.63	0.78	1.24	0.63	0.59	1.39	1.39
RMSE of MAP estimates (%)	0.44	0.43	0.44	0.43	0.21	0.44	0.43	0.21	1.05	1.18
Improvement Factor (row 1/row 2)	1.3	1.4	3.0	1.5	3.7	2.8	1.5	2.8	1.3	1.2

STD. (This is equivalent to RMSE= $\sqrt{\left\langle(\hat{\mathbf{m}}-\mathbf{m}_{true})^2\right\rangle}$, where the averaging is over 10^6 realizations of noise.)

This process is repeated for MAP estimates, and the results are compared. For both estimation methods, and for all ten parameters, the bias is less than 0.01% of the true parameter values (that is, 1 part in 10,000). The STD is therefore the same as the RMSE, to four significant digits. In the first two rows of Table 4.2 we report the RMSE for each method and for each of the ten parameters, as percentages of the true parameter values. In the third row we report the factor by which the RMSE of MAP estimates is lower than the RMSE of algebraic estimates[8].

To summarize our results with a single number, we take the average of these ten improvement factors. It is 2.04 – that is, the RMSE of MAP estimates is about two times smaller than the RMSE of algebraic estimates.

To establish the accuracy of this number, the entire procedure of the last four paragraphs was repeated 7 times to provide 7 estimates of the average improvement factor. The average of these 7 numbers was 2.041, with a STD of 0.001. Also, the results reported in Table 4.2 were the same (to two decimal places) in all 7 cases.

The cost for the increased accuracy of MAP estimates is an increase in computation. On average, a MAP estimate requires 40,000 times more computation than an algebraic estimate (this could be reduced if pains were taken to increase the efficiency of the eigendecomposition, etc.). But MAP estimation is still quite

---

[8] A likely explanation for the assymetry in the RMSE of MAP estimates of T_1 versus those of T_2 is as follows. The variance of n_5 is smaller than that of n_6 because $T_C < T_H$. Since n_5 is added to T_1 while n_6 is added to T_2, we end up with less uncertainty in T_1 than in T_2.

tractable: as of this writing, a standard computer can perform a MAP estimate in about 0.06 seconds, even without any optimization of the eigendecomposition[9].

Improvement as a Function of $\mathbf{m}_{true}$

The improvement in accuracy of MAP estimates (over algebraic estimates) is a weak function of $\mathbf{m}_{true}$. If we repeat the above procedure for 100 different $\mathbf{m}_{true}$, chosen randomly within expected ranges[10] , we find that the RMSE of MAP estimates is consistently the same as the values in Table 4.2. The RMSE of algebraic estimates is slightly lower than the values in Table 4.2, resulting in an average improvement factor that ranges from 1.86 to 2.03, with the mean being 1.90.

4.6 Sampling the Posterior Pdf

We now turn to exploration of the more complete answer to the calibration problem, the posterior pdf $p(\mathbf{m}|\mathbf{v})$. If samples of $p(\mathbf{m}|\mathbf{v})$ are available, they can be used to visualize the posterior pdf, to report it, or to calculate its mean, $\mathbf{m}_{MMSE}$. In this section we describe how to generate such samples and then use them for the aforementioned purposes.

4.6.1 Sampling $p(\mathbf{m}|\mathbf{v})$ by the Rejection Method

Samples of the posterior pdf, $p(\mathbf{m}|\mathbf{v})$, can be generated by the well known rejection method (a.k.a. acceptance-rejection sampling) [3] [27]. First, samples of

---

[9] The above results were obtained using MATLAB's fminsearch's default options: tolFun= 10^{-4} and tolX= 10^{-4} (and with the default maxFunEvals = maxIter= 2000, these tolerances are consistently achieved). When the options are changed in order to produce convergence to a more precise peak (specifically, tolFun= 10^{-11}, tolX= 10^{-11}, maxFunEvals = 10^5, and maxIter= 10^4), the average improvement in RMSE is a factor of 2.042 (for four runs of 10^6 estimates each) and the increase in computation is 77,000. This demonstrates that the default precision is adequate.

[10] c_v, c_h, c_p, and c_m are chosen independently from normal distributions with mean of 450 and STD of 17 mV/mW. G_1 is chosen from a normal distribution with mean of 18e6 and STD of 2.7e6 W/W; $10log_{10}(g)$ is chosen from a normal distribution with mean of 0 and STD of 1 dB; s is chosen from a normal distribution with mean of $1/\sqrt{2}$ and STD of $0.02/\sqrt{2}$; α_e is chosen from a normal distribution with mean of 0.934 and STD of 0.01; and T_1 and T_2 are chosen independently from normal distributions with mean of 310 and STD of 1 K. Also, T_C is chosen from a normal distribution with mean of 288 and STD of 0.5 K while T_H and T_{CN} are chosen independently from normal distributions with mean of 800 and STD of 2 K.

G_{vv}, G_{hh} and G_{pU} are proposed from independent, uniform distributions. These distributions are centered on an initial guess such as the MAP estimate.

With the proposed G_{vv}, G_{hh} and G_{pU} coordinates, we next find proposed G_{ph}, G_{mh}, G_{pv}, G_{mv}, and G_{mU} coordinates from the constraint equations, (4.21) through (4.23). Also, T_1 and T_2 coordinates are proposed from independent, uniform distributions centered on the initial guess. Each set of ten proposed coordinates now comprises one proposed sample of $p(\mathbf{m}|\mathbf{v})$, which we denote $\mathbf{m}_{prop}$.

Because the constraint equations were used, each $\mathbf{m}_{prop}$ is a sample from a uniform distribution over a region of the constraint manifold. Let the constant value of this uniform distribution be denoted by k. In order to correctly generate samples of $p(\mathbf{m}|\mathbf{v})$, we must accept each $\mathbf{m}_{prop}$ with probability P where

$$P = \frac{p(\mathbf{m}_{prop}|\mathbf{v})/k}{max_{\mathbf{m}}\left(p(\mathbf{m}|\mathbf{v})/k\right)} = \frac{p(\mathbf{m}_{prop}|\mathbf{v})}{max_{\mathbf{m}}\left(p(\mathbf{m}|\mathbf{v})\right)}. \tag{4.29}$$

The numerator of (4.29) is readily calculated using (4.24), with V_1 found from a numerical eigendecomposition of C and C calculated from $\mathbf{m}_{prop}$. The denominator of (4.29) is the peak value of $p(\mathbf{m}|\mathbf{v})$. The search for this peak is the same search that finds the MAP estimate of $\mathbf{m}$ – that is, the denominator is simply $p(\mathbf{m}_{MAP}|\mathbf{v})$.

In the rejection method just described, samples are proposed over only a rectangular region (determined by the upper and lower bounds of the uniform 1-D distributions) of the constraint manifold, rather than over the entire constraint manifold (which has infinite extent). This effectively truncates $p(\mathbf{m}|\mathbf{v})$ to the rectangular region. This is equivalent to using a 10-D uniform pdf $p(\mathbf{m})$ in (4.8) that truncates to the rectangular region. In other words, the above method adds external information (even if it is false) that the true $\mathbf{m}$ definitely lies in the rectangular region. This is a practical necessity when using uniform distributions to propose samples, since the probability of accepting a proposed sample goes to zero as the rectangular region is enlarged in order to better cover the infinite extent of the constraint manifold. The truncation is insignificant in the current problem however, see Section 4.6.2[11].

The bounds of the five 1-D uniform distributions of G_{vv}, G_{hh}, G_{pU}, T_1 and T_2 are somewhat arbitrary. Looser bounds cause less truncation but decrease the rate at which proposed samples are accepted. A simple method of choosing the bounds is to take the STD (= RMSE since bias ≈ 0) of simulated algebraic estimates or MAP estimates, given in Table 4.2, and let the bounds be a certain multiple of STD above and below the MAP estimates.

Numerical note: the sampling process is *much* faster when all the G_{xx} (or voltages) are scaled by an appropriate factor. A factor of 10^7 was used in generating the Figures.

4.6.2 Using the Samples

Marginal Posterior Pdfs

Once we have a number of samples of $p(\mathbf{m}|\mathbf{v})$, we can immediately obtain plots of the marginal (meaning one-dimensional) probability distribution for the ith parameter by simply binning the ith coordinate values of the samples (we believe that this follows from [3]). Some marginal pdfs obtained in this manner are shown in Fig. 4.1.

Fig. 4.1 demonstrates that the marginal pdfs are symmetric[12]. This verifies the claim made earlier that MAP estimates (mode of marginal pdfs) are the same as MMSE estimates (means of marginal pdfs). Furthermore, when MMSE estimates are made, they have the same error statistics as MAP estimates, as given in Table 4.2 and Section 5.4.

Fig. 4.1 also demonstrates that the marginal pdfs are Gaussian (or at least very nearly so). The Gaussians that are plotted have STD from the second line of

---

[11] If Gaussian distributions were used to propose samples, then no truncation occurs but (4.29) is more difficult to evaluate. Using Gaussians may be only slightly more difficult, and if successfully implemented, it would also result in a *much* higher sample acceptance rate.

[12] Some assymetry appears in Fig. 4.1, but this assymetry is a consequence of the finite number of samples used to generate the marginal pdfs. This assymetry diasappears as the number of samples increases.

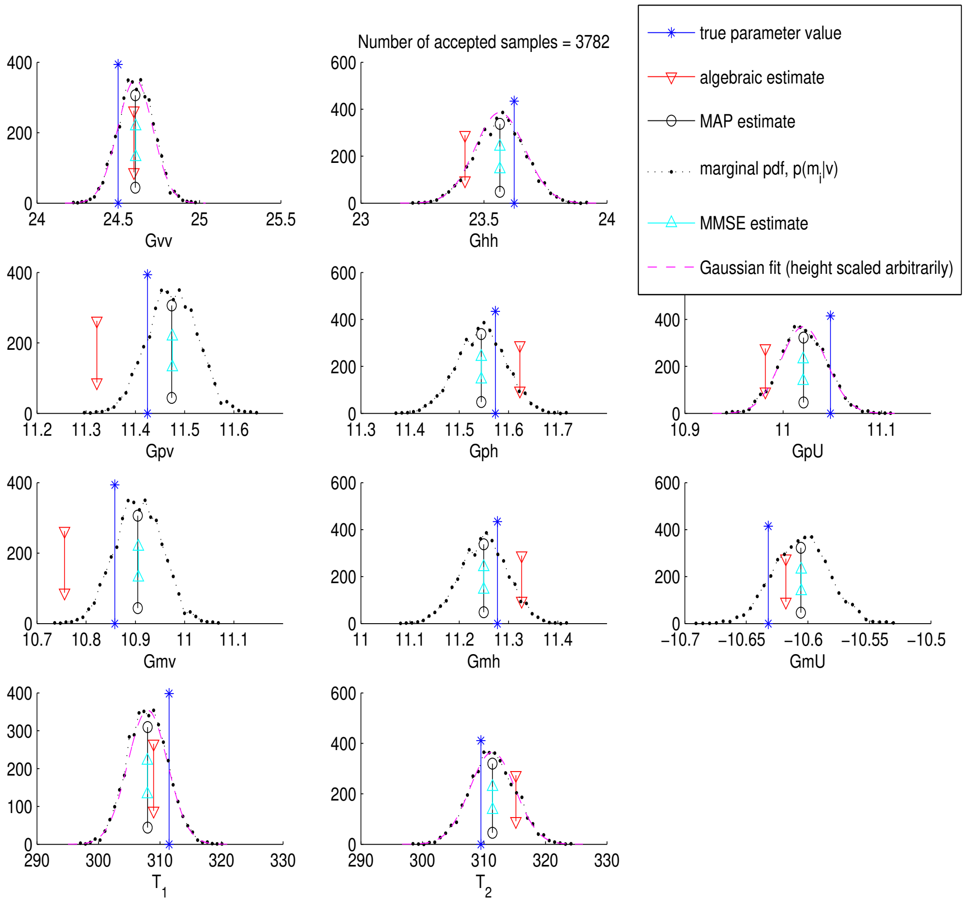

Figure 4.1: Marginal pdfs for the eight radiometer gains and two noise temperatures, along with various estimates of the parameters. The height scale is arbitrary, being the number of samples in each bin. As we use more accepted samples to generate the empirical pdfs, the heights increase and the empirical pdf matches the Gaussian more closely.

Table 4.2. Hence, each marginal pdf can be completely characterized by its mode (the MAP estimate) and STD (= RMSE of MAP estimates, given in Table 4.2).

Additional Information in Joint Pdfs

Joint pdfs can convey many times more information than the marginal pdfs that are extracted from them. For example, consider the two variables G_{vv} and T_1, whose joint posterior pdf[13] is depicted in Fig. 4.2. The density of dots (samples of the pdf) illustrates the probability density. This figure is made from the same samples as Fig. 4.1.

The 2-D joint pdf in Fig. 4.2 contains significant correlation information. For example, it shows that there is a fair chance that $G_{vv} \approx 24.5$ and $T_1 \approx 311$, but almost no chance that $G_{vv} \approx 24.5$ and $T_1 \approx 304$. If we had reported only the 1-D marginal pdfs in Fig. 4.1, however, both possibilities would have been reported as equally likely.

The current practice of reporting only the mean and variance for each parameter is equivalent to reporting 1-D marginal Gaussian pdfs for the parameters [3]. The only joint pdf that can logically be reconstructed from marginal pdfs is the product of the marginal pdfs (this follows from [3]). An example of such a reconstruction is shown in Fig. 4.3. All correlation information is lost, as well as any non-Gaussian characteristics of the posterior pdf.

The situation is even more pronounced for the parameters in the current problem whose estimation is 100% correlated (due to the constraint equations). For example, a 2-D joint pdf of G_{vv} and G_{pv} is shown in Fig. 4.4 (also made from the same samples as Fig. 4.1 and Fig. 4.2). The pdf is completely concentrated along a 1-D line in the 2-D space. If only means and variances were reported, the appearance would be similar to Fig. 4.3.

We have demonstrated that the most complete answer to an estimation problem is a joint pdf, but how can a 10-D posterior pdf be reported? We can report the equation for it, such as (4.24). A numerical alternative is to simply report a

---

[13] Given a particular set of voltages measured by a radiometer.

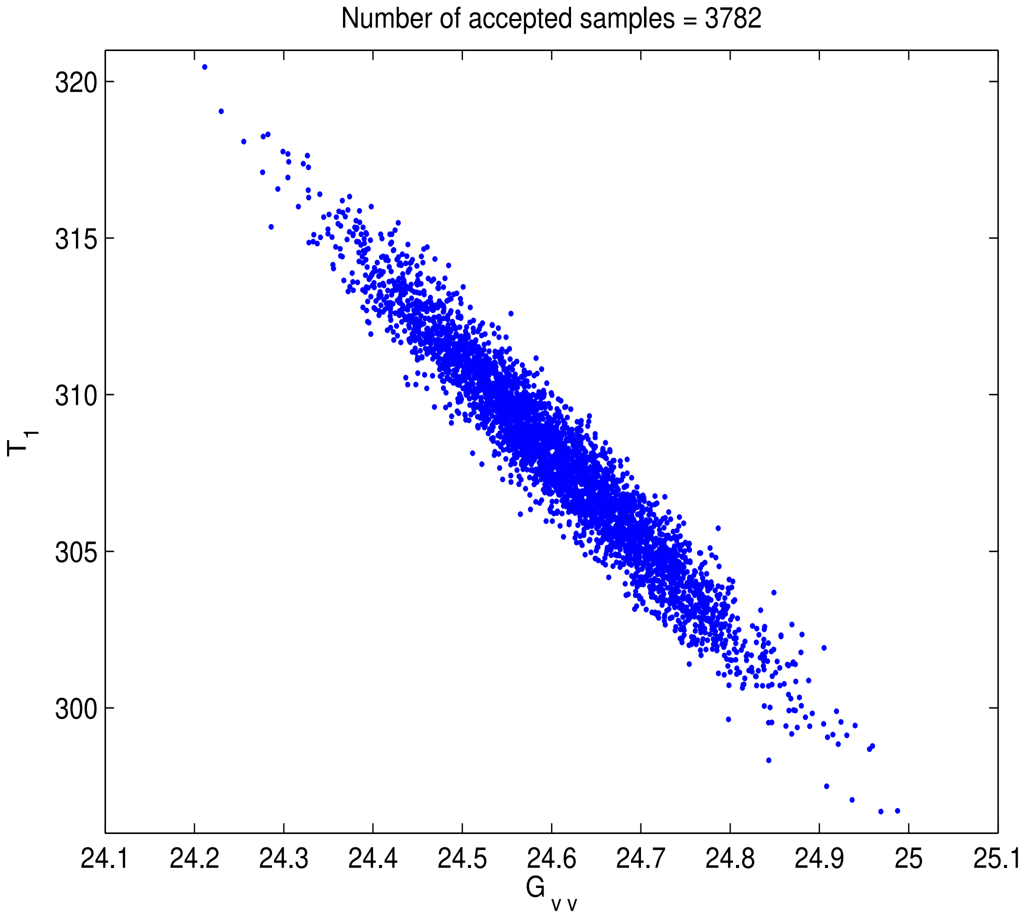

Figure 4.2: Dots showing the G_{vv} and T_1 coordinates of samples of $p(\mathbf{m}|\mathbf{v})$. The density of dots illustrates the 2-D joint pdf $p(G_{vv}, T_1|\mathbf{v})$. The samples are from the same simulation that produce Fig. 4.1.

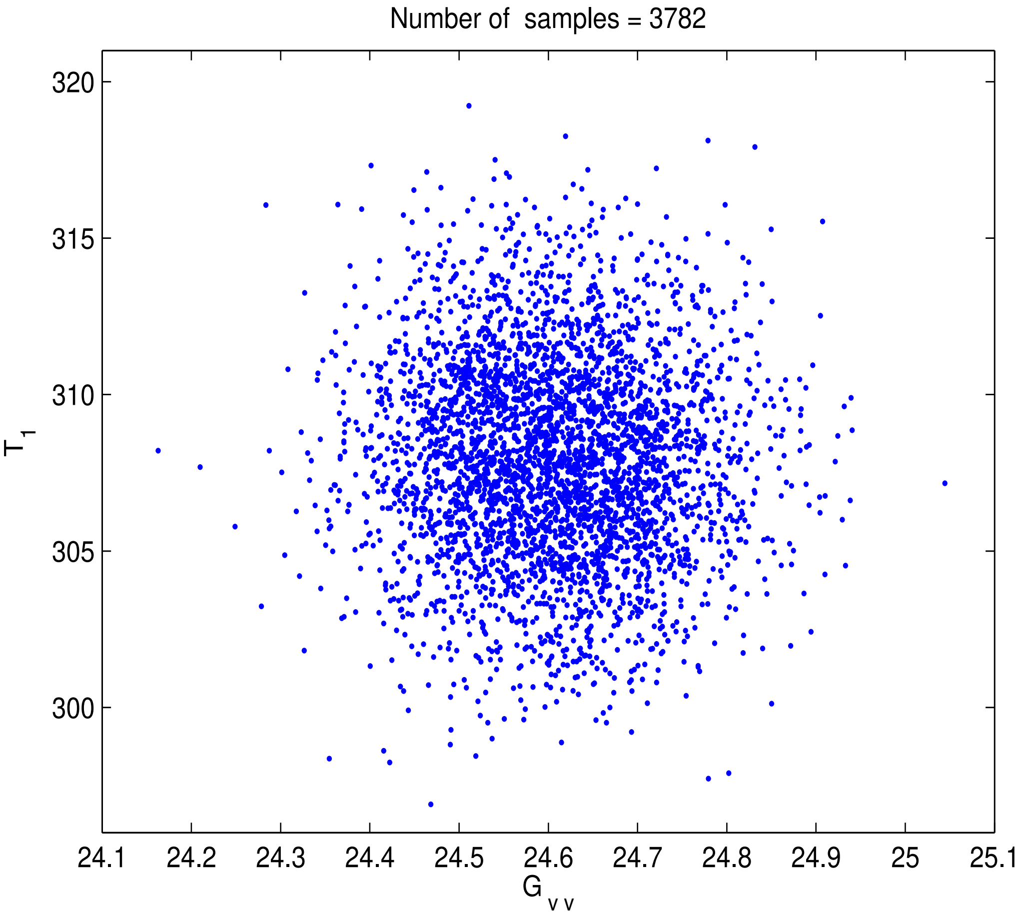

Figure 4.3: Depiction of the 2-D joint pdf for G_{vv} and T_1 which conveyed by reporting only means and variances – using means and variances of samples in Fig. 4.2. Comparison with Fig. 4.2 shows the loss of information which occurs by reporting only means and variances.

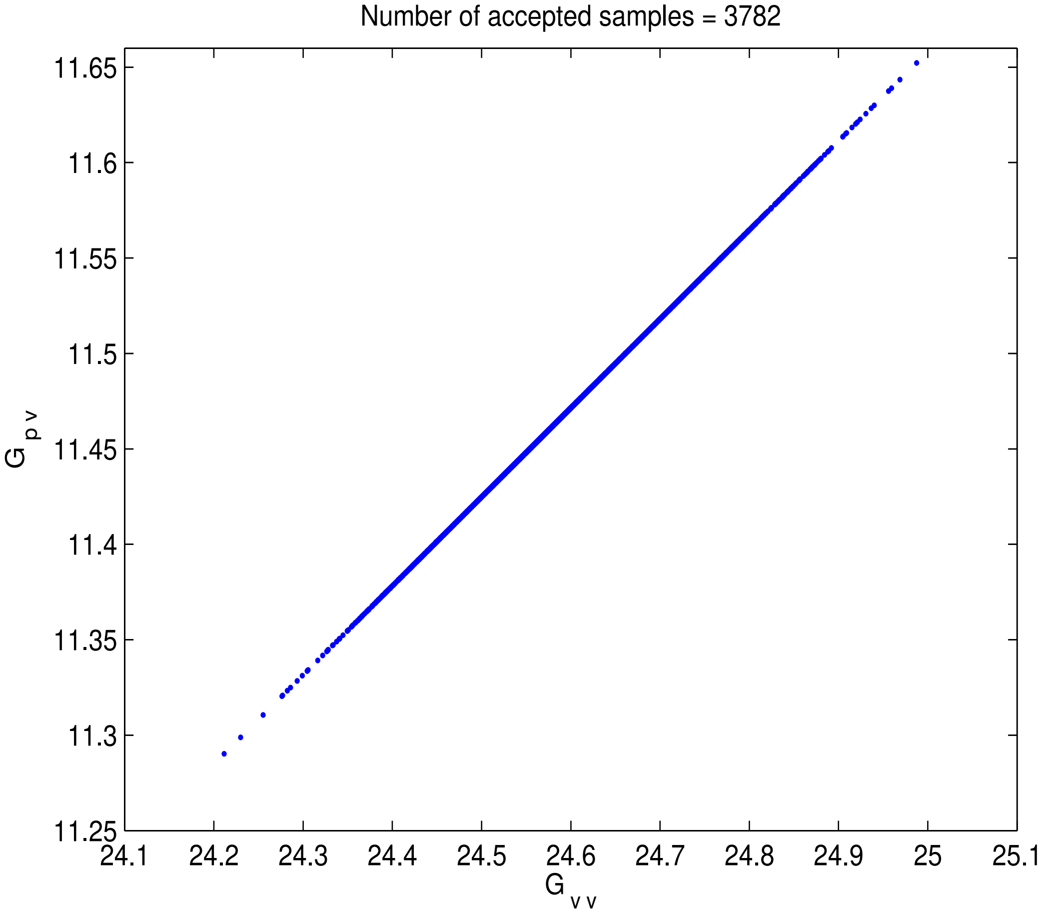

Figure 4.4: Dots showing the G_{vv} and G_{pv} coordinates of samples of $p(\mathbf{m}|\mathbf{v})$. The density of dots illustrates the 2-D joint pdf $p(G_{vv}, G_{pv}|\mathbf{v})$.

large number of samples of the pdf because most, if not all, calculations done with a posterior pdf can be done using these samples [3]. Finally, a very succint alternative is available when a multidimensional pdf is sufficiently Gaussian, as in the present case. In such cases, all of the information can be conveyed by a vector of means and a covariance matrix. For example, the correlation matrix for samples of $p(\mathbf{m}|\mathbf{v})$, calculated from samples like those displayed in Fig. 4.1 and Fig. 4.2, is

$$\begin{bmatrix} 1.00 & 0.04 & 1.00 & 0.04 & 0.22 & 1.00 & 0.04 & -0.22 & -0.96 & -0.02 \\ 0.04 & 1.00 & 0.04 & 1.00 & 0.13 & 0.04 & 1.00 & -0.13 & -0.03 & -0.96 \\ 1.00 & 0.04 & 1.00 & 0.04 & 0.22 & 1.00 & 0.04 & -0.22 & -0.96 & -0.02 \\ 0.04 & 1.00 & 0.04 & 1.00 & 0.13 & 0.04 & 1.00 & -0.13 & -0.03 & -0.96 \\ 0.22 & 0.13 & 0.22 & 0.13 & 1.00 & 0.22 & 0.13 & -1.00 & -0.17 & -0.08 \\ 1.00 & 0.04 & 1.00 & 0.04 & 0.22 & 1.00 & 0.04 & -0.22 & -0.96 & -0.02 \\ 0.04 & 1.00 & 0.04 & 1.00 & 0.13 & 0.04 & 1.00 & -0.13 & -0.03 & -0.96 \\ -0.22 & -0.13 & -0.22 & -0.13 & -1.00 & -0.22 & -0.13 & 1.00 & 0.17 & 0.08 \\ -0.96 & -0.03 & -0.96 & -0.03 & -0.17 & -0.96 & -0.03 & 0.17 & 1.00 & 0.02 \\ -0.02 & -0.96 & -0.02 & -0.96 & -0.08 & -0.02 & -0.96 & 0.08 & 0.02 & 1.00 \end{bmatrix} \tag{4.30}$$

This provides the important information found in Fig. 4.2 and Fig. 4.4 but not in Fig. 4.1 or Fig. 4.3 – for example, that the 2-D pdf for G_{vv} and T_1 has a correlation coefficient of -0.96.

As a final note: the relative amount of information preserved by various joint pdfs can be compared quantitatively by calculating the entropy of such pdfs. For example, the joint pdf in equation (4.24) has much less entropy than a joint pdf which is reconstructed from marginal pdfs for the parameters.

MMSE Estimation

Another use for the samples of a distribution is in calculating its expected value. As mentioned above, this expected value (mean) is the MMSE estimate of $\mathbf{m}$,

$$\mathbf{m}_{MMSE} = E_{\mathbf{m}}(p(\mathbf{m}|\mathbf{v})) = \begin{bmatrix} \int_M \; m_1 p(\mathbf{m}|\mathbf{v}) \; d\mathbf{m} \\ \int_M \; m_2 p(\mathbf{m}|\mathbf{v}) \; d\mathbf{m} \\ \cdot \\ \cdot \\ \cdot \\ \int_M \; m_{10} p(\mathbf{m}|\mathbf{v}) \; d\mathbf{m} \end{bmatrix} \tag{4.31}$$

where M is the 10-D space of the ten calibration parameters. For example, to estimate G_{vv},

$$m_{1,MMSE} = \int_M m_1 p(\mathbf{m}|\mathbf{v}) \; d\mathbf{m} \tag{4.32}$$

and similarly for the other nine parameters. Equation (4.32) could be evaluated using a quadrature rule. But it can also be numerically implemented by generating samples of $p(\mathbf{m}|\mathbf{v})$ and then simply taking the mean of the m_1 coordinate of those samples. As the number of samples increases, this procedure converges to (4.32) [3].

4.7 Information on Hardware Gain Parameters

As a tangential but useful extension of the previous results in this chapter, we can use the probabilistic method developed above to obtain information on the radiometer hardware parameters that comprise the eight gains G_{xx}. The definition of these gains in terms of hardware paramters, found by comparing equations (1) and (17) in [2] with one another, is reproduced here:

$$\begin{bmatrix} G_{vv} & 0 & 0 \\ 0 & G_{hh} & 0 \\ G_{pv} & G_{ph} & G_{pU} \\ G_{mv} & G_{mh} & G_{mU} \end{bmatrix} \equiv kB \begin{bmatrix} c_v G_1 & 0 & 0 \\ 0 & c_h G_2 & 0 \\ c_p s^2 G_1 & c_p(1-s^2)G_2 & c_p s\sqrt{1-s^2}\alpha_e\sqrt{G_1 G_2} \\ c_m(1-s^2)G_1 & c_m s^2 G_2 & -c_m s\sqrt{1-s^2}\alpha_e\sqrt{G_1 G_2} \end{bmatrix}. \tag{4.33}$$

4.7.1 Analytical Results

By replacing the eight G_{xx} in (4.24)[14] with their component definitions on the right side of (4.33), we immediately obtain the joint posterior pdf for the eight hardware gain parameters (G_1, G_2, α_e, c_v, c_h, c_p, c_m, and s) plus T_1 and T_2. (The bandwidth B could also be considered a parameter, but for this work we treat it as a known constant. The parameter k is Boltzmann's constant.) For brevity, we simply summarize our discoveries about this pdf. Details of the derivations are given in Appendix D.

First, the transformed constraint equations dictate that the parameter s is uniquely determined by the measured calibration voltages:

$$s = \sqrt{\frac{AD + \sqrt{ADBC}}{AD - BC}}, \tag{4.34}$$

where

$$\begin{aligned} A &\equiv v_{p,C}v_{h,H} - v_{h,C}v_{p,H}, \\ D &\equiv v_{m,C}v_{v,H} - v_{v,C}v_{m,H}, \\ B &\equiv v_{p,C}v_{v,H} - v_{v,C}v_{p,H}, \\ C &\equiv v_{m,C}v_{h,H} - v_{h,C}v_{m,H}. \end{aligned}$$

It is worth noting that this closed-form solution for s fell out of the equations developed for a probabilistic estimation of the parameters. If a simple algebraic approach had been attempted, the closed-form solution may not have been found, as (4.34) is a rather complicated function of the measured voltages.

Another remarkable fact is that there is no uncertainty in this estimate for s – it is not affected by $NE\Delta T$. (In simulation therefore, this estimate is the exact value of the true s; but in reality, any uncertainties not captured by our model in (4.10), such as imperfect knowledge of T_C and T_H, will cause this estimate to have some error.)

---

[14] The G_{xx} are found both in $\mathbf{m}$ and in the constraint equations.

The constraint equations also dictate that c_h, c_p, and c_m are constrained as follows:

$$\begin{aligned} c_h &= c_v\sqrt{\frac{AC}{BD}}, && (4.35) \\ c_p &= c_v\frac{\sqrt{A}}{E\sqrt{D}}\left(\sqrt{AD}-\sqrt{BC}\right), && (4.36) \\ c_m &= -c_v\frac{\sqrt{C}}{E\sqrt{B}}\left(\sqrt{AD}-\sqrt{BC}\right), && (4.37) \end{aligned}$$

where $E \equiv v_{v,C}v_{h,H} - v_{h,C}v_{v,H}$. One interpretation of these equations is that the *ratio* of any two c_x is perfectly resolved – it is equal to a function of the measured voltages.

Finally, the remaining four hardware gain parameters (G_1, G_2, α_e, and c_v[15]) are unaffected by the constraint equations.

4.7.2 Numerical Results

Additional information on the hardware gain parameters is readily obtained by examining samples[16] of the posterior pdf for them. In the first place, the s coordinate of the samples has no variance, confirming the analytical conclusion above that estimation of s is unaffected by $NE\Delta T$.

The samples can readily be used to plot pdfs for pairs of the hardware parameters, as shown in Fig. 4.5. These plots reveal several characteristics.

First, an intuitive "conservation of information" principle would suggest that the ability to perfectly resolve s must be compensated by a lack of ability to resolve other parameters. This is indeed the case: As shown by Fig. 4.5, c_v, G_1, and G_2 cannot be individually resolved. We can only resolve their pairwise products[17]. Marginal pdfs for c_v, G_1, and G_2 are essentially uniform[18].

---

[15] Note that (4.35) through (4.37) can be rearranged to put any three of the c_x in terms of the remaining one; thus any one of the c_x can be considered unconstrained with the remaining three tightly constrained via such equations. c_v has been chosen arbitrarily to be considered the unconstrained parameter in this work.

[16] The method for producing these samples is very similar to the method described in Section 4.6.

[17] For example, c_vG_1 is resolvable – as shown by (4.33), it is $(kB)^{-1}$ times G_{vv}, and G_{vv} is well resolved by either the algebraic method or probabilistic method discussed earlier in this chapter.

[18] These uniform marginal pdfs are only limited by prior knowledge. However, if we have tighter prior knowledge on one parameter, say on c_v, then good resolution of a product such as c_vG_1 can cause tighter bounds on the other parameter, G_1. This effect can be seen in several subplots of Fig. 4.5.

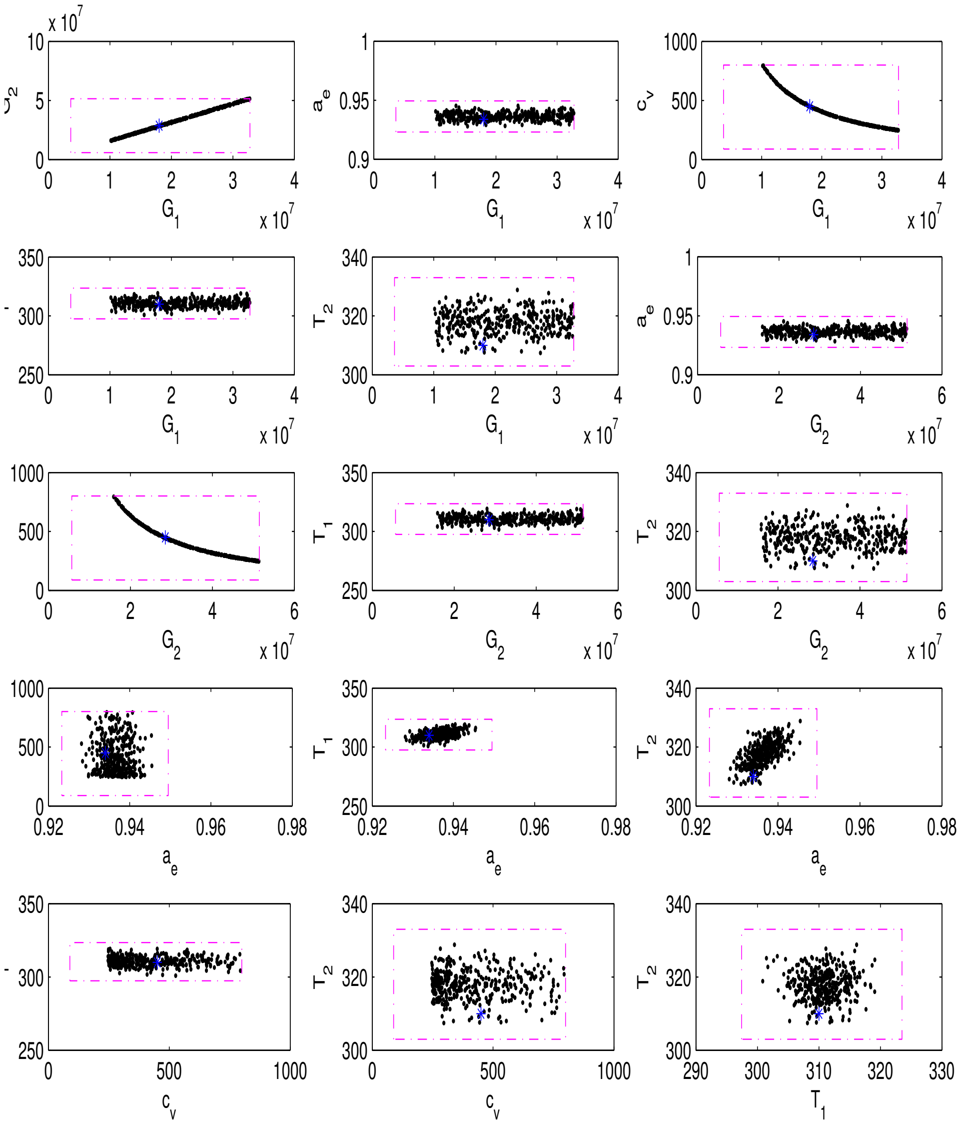

Figure 4.5: Two-dimensional pdfs for the four unconstrained radiometer hardware gain parameters (G_1, G_2, α_e, and c_v) and two noise temperatures (T_1 and T_2). The density of dots illustrates the pdfs, whereas asterisks show the location of the true parameter values. The bounding boxes show the bounds of the uniform pdfs from which samples were proposed. These plots illustrate that it is only possible to obtain useful estimates of the hardware gain parameter α_e (in addition to s which is not illustrated).

A second observation is that α_e is resolved well. The average RMSE of MAP estimates of α_e is 0.33 %, whether using the typical values given in Table 4.1 or randomized values for the true hardware parameters.

4.8 Conclusion

This chapter demonstrates the advantages (and drawbacks) of estimating calibration parameters via a probabilistic approach rather than a conventional algebraic approach. The first advantage of the probabilistic approach is a reduction in error. By exploiting statistical knowledge of measurement noise using Bayesian estimation, the RMSE of parameter estimates is reduced by a factor of two compared to estimation without such knowledge.

This work also illustrates the principle that much more information can be conveyed by a probability distribution for a set of parameters than by simple estimates comprised only of marginal means and variances. In particular, valuable covariance information is conserved via multidimensional pdfs. The generation and utility of samples of such pdfs has been demonstrated.

Finally, this work shows that a probabilistic approach reveals valuable information on the eight hardware parameters that comprise the overall channel gains in the class of radiometer which is analyzed in this work. The probabilistic approach provides accurate solutions for two of these parameters. An algebraic approach to such solutions would be difficult if not intractable. For the other six parameters, the probabilistic approach demonstrates that only products or ratios of pairs of the parameters can be resolved.All of these results are valuable for diagnosis of radiometer anomalies.

Most of the principles employed herein are well known in estimation theory. However, this work is their first published application to microwave radiometer calibration.

Chapter 5

Adaptive Inference of Polar Air Temperatures

In the previous two chapters, probability theory was used to improve the analysis and usage of microwave radiometer data for situations which are well modeled by a forward model. This chapter considers a problem for which a solid forward model does not exist and an empirical inversion is therefore advantageous.

The previous empirical solution to the problem involved fitting a sinusoidal curve to training data and using this curve to estimate desired parameters from available measurements. As in previous chapters, a better solution to the inverse problem is possible through a probabilistic approach. This approach is to use training data to form an empirical pdf which relates measurements to the desired parameter. This solution is not only more accurate than the curve-fitting solution (especially as the amount of training data increases) but is also significantly more robust because an empirical pdf is more adaptive or flexible than a simple curve. This additional adaptability enables the pdf to capture an additional physical phenomenon.

5.1 Introduction

Antarctica and Greenland represent 11% of the Earth's land surfaces, so measuring temperature there is of some importance to weather forecasting and climate prediction. However, weather stations to monitor such areas are quite sparse in both time and spatial distribution due to the high costs of deploying and maintaining stations on barren ice sheets. Consequently, the feasibility of retrieving surface temperature from satellite measurements has been investigated by several groups [28]

[29]. Polar orbiting satellites are well suited for this since their orbit geometry permits them to measure both polar regions many times each day.

Satellite records can also be very dependable. For example, the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) has provided a measurement record that is essentially unbroken since 20 Sep 2002. In contrast, weather station records for the interior of Antarctica are plagued with gaps (with the exception of south pole stations, which are not considered in this work because orbit geometry prevents most satellites from observing the poles). In this chapter AMSR-E data is used to retrieve near-surface air temperature in Antarctica.

The state of the art in retrieving near-surface air temperature T from microwave emissions is a technique described by Shuman et al. [29]. This method, along with infrared methods, is summarized in Section 5.2. In Section 5.3, we expound an adaptive method with probability theoretic underpinnings which provides more accurate estimation of T from microwave emissions. Performance of the two methods is compared in Sections 5.4 and 5.5. Conclusions and future work are discussed in Section 5.6.

5.2 Previous Methods

In this section we review previous methods of retrieving air temperature over ice sheets from satellite measurements. These methods can be classified as either infrared or microwave.

Satellite measurements of infrared emissions can be used to retrieve surface temperature quite accurately under clear sky conditions [30, p. 6-7]. Unfortunately, this method cannot retrieve temperature where clouds are present, and clouds typically cover a major portion of Antarctica. Furthermore, good cloud discrimination cannot be achieved during darkness. The combination of these two effects undermines infrared temperature retrieval for more than half of every year.

Although satellite measurements of microwave emissions have poorer resolution than infrared measurements, at certain frequencies such as 37 GHz they are largely unaffected by atmospheric conditions [6, p. 20-21] or darkness. Shuman et al. [29] compare ground truth T with 37 GHz vertically polarized (v-pol) brightness temperature B as measured by the SSM/I instruments. They find that the familiar relationship

$$B = \epsilon T \tag{5.1}$$

can be used to find a reasonably accurate estimate of T from B if the effective emissivity ϵ of the surface is modeled as a sinusoidal function of day of year (DOY), denoted $\epsilon_{sinus}(DOY)$, with a period of one year.

The mean, amplitude, and phase of Shuman et al.'s[29] sinusoidal emissivity ϵ_{sinus} are determined empirically as demonstrated in Fig. 5.1. A time series of B data (black in upper plot) is divided by a coincident T time series (blue in upper plot), yielding an empirical emissivity time series (red in lower plot). The particular sinusoid which minimizes the root-mean-square error (RMSE) between the time series and the sinusoid becomes ϵ_{sinus}, as shown by the dashed line in the lower plot of Fig. 5.1. With this emissivity $\epsilon_{sinus}(DOY)$ in hand, estimation of T at times when only B and DOY are available is

$$\hat{T} = B/\epsilon_{sinus}(DOY). \tag{5.2}$$

Application of this method in Greenland by Shuman et al. shows that this method achieves a good estimate of smoothed daily temperatures. In Section 5.4 we apply this method to three Antarctic sites and obtain similar results. Hereafter we refer to this method as the sinusoidal method.

Shuman et al. suggest that the cycle in empirical emissivity is due to a seasonal cycle in the actual emissivity of the ice. While this may be true in part, the cycle may also be a hysteresis effect caused by heat flow to and from deeper ice layers. Since deeper layers have a more constant temperature than the surface, heat flows upwards during the cooling season (roughly February through May). This causes the surface ice to be warmer than the air, which is manifested as an apparent increase in emissivity. In the warming season (roughly August through November), the reverse occurs. The possibility of this effect is consistent with heat flow comments in the introduction of Surdyk [28], in which it is also noted that a pioneering study using a

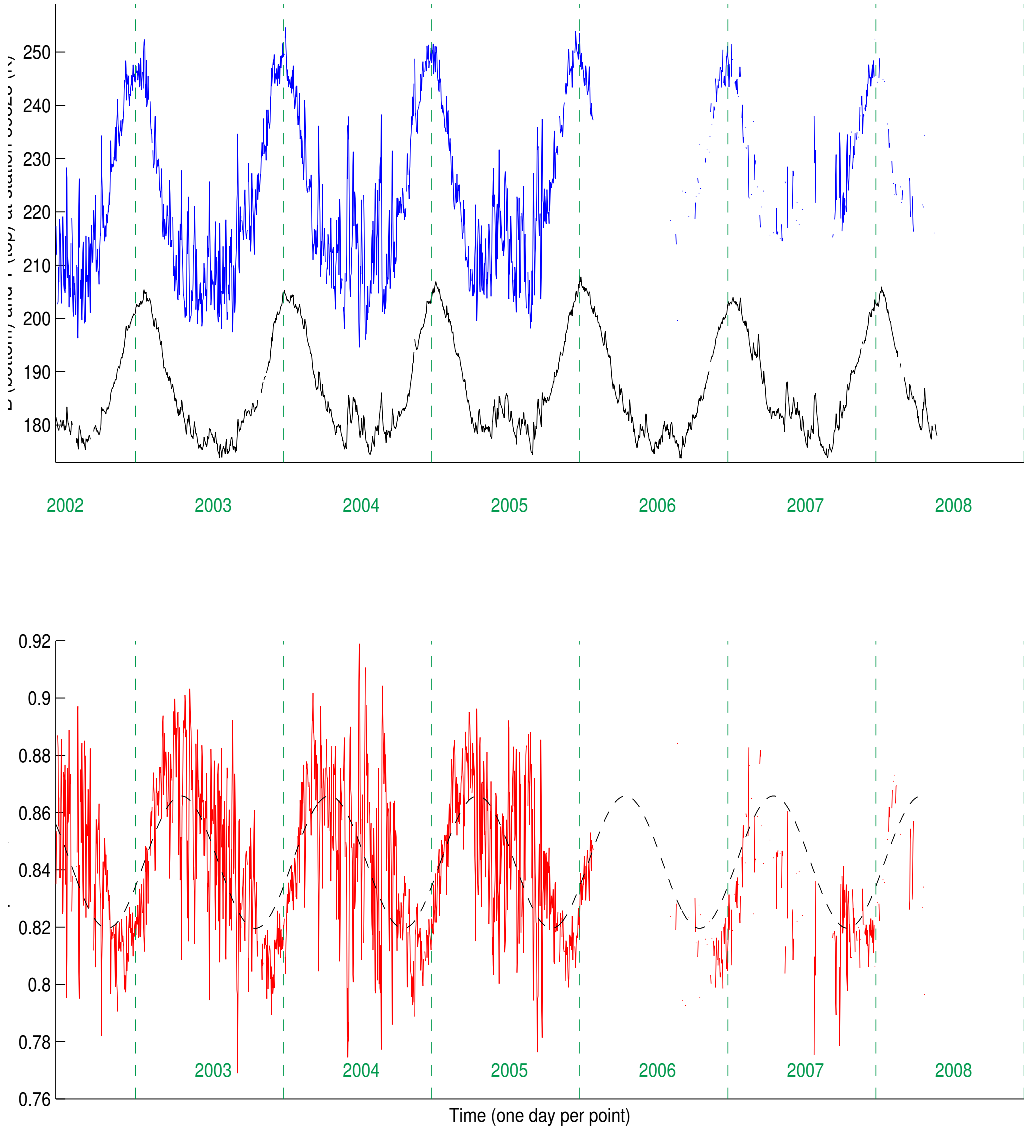

Figure 5.1: Time series of coincident (blue in upper plot) T, (black in upper plot) B, and (red in lower plot) empirical ϵ measurements at Antarctic automatic weather station 89828 (Dome C), along with the (dashed in lower plot) sinusoidal fit $\epsilon_{sinus}(DOY)$ which minimizes RMSE. Note the gaps in the T record indicating that the automatic weather station is faltering.

radiative transfer model by Zwally [31] indicates that the annual emissivity is nearly constant.

5.3 Pdf Method

We desire to use satellite measurements of microwave earth emissions B, together with DOY information, to obtain the best possible estimates of near-surface air temperature T. We concur with Shuman et al. in the conclusion that formulating a geophysically based forward model of B from T and then inverting it to obtain T from B is very difficult because of the multitude of parameters in such a model which are difficult to know (such as ice grain size and geometry, layer thicknesses, heat flow in the firn, and thermal effects of wind pumping), although this approach is considered in [28]. We instead seek to empirically determine the relationship between T, B, and DOY, using coincident records of these three variables.

The sinusoidal emissivity model discussed in Section 5.2 is reasonable. But if an empirical relationship is to be derived, why should it be constrained to a sinusoidal function? The simplicity of a sinusoid is attractive, but a more flexible relationship achieves greater accuracy. Deriving a more accurate relationship is the task accomplished in this section. We refer to this method as the probability distribution function (pdf) method, for reasons that become apparent.

5.3.1 Estimating T from B only

We begin by demonstrating how the pdf method estimates T from only B information. Consider the three years of coincident T and B data for an Antarctic weather station which are plotted in Fig. 5.2. The coordinates of each point are the average T for a day and the average B for the same day (see Appendix E for details on the datasets used). This plot shows the empirical information which we have about the relationship between T and B at this location.

An intuitive description of the proposed pdf method is as follows. For a particular satellite measurement denoted $\bar{B}$, such as $\bar{B} = 180$ K, the density of the points in Fig. 5.2 along the line $B = 180$ can be viewed as $p(T|\bar{B} = 180)$, a probability distribution function for T given the knowledge that $\bar{B} = 180$. This pdf itself is the most complete information on T, given the input $\bar{B}$ and the empirical relationship depicted in Fig. 5.2. Numerical estimates of T can be chosen based on $p(T|\bar{B})$: choosing the mean of $p(T|\bar{B})$ provides the minimum mean squared error (MMSE) estimate of T, given $p(T|\bar{B})$ [20, p. 175].

In more precise terms, the proposed method solves an inverse problem (obtaining T from B) using probability theory. We follow the exposition of Tarantola [3, Ch. 1], which is similar to Bayes' theorem but more general and complete. The points plotted in Fig. 5.2 can be considered to be samples of our best available estimate of the joint theoretical pdf of B and T, which we denote $\Theta_1(b,t)$ (see [3, section 1.3]). This pdf captures our empirical knowledge of the relationship between B and T.[1] Satellite measurement information, together with its uncertainty, is captured in another pdf, denoted by $\rho_B(b)$ (see [3, section 1.4.1]). We model this pdf with a Gaussian pdf whose mean is a particular satellite measurement $\bar{B}$ and whose standard deviation (STD) is σ, that is, $\rho_B(b) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(b-\bar{B})^2/(2\sigma^2)}$. Then[2] the output of the pdf method is a marginal pdf for T, denoted $\sigma_T(t)$, that is found as follows:

$$\sigma_T(t) \quad = \quad k \int \rho_B(b)\ \Theta_1(b,t)\ db, \tag{5.3}$$

where k is an unknown constant and the integration is carried out over all values of b.

In words, we multiply the joint distribution $\Theta_1(b,t)$ depicted in Fig. 5.2 by the measurement information $\rho_B(b)$, so as to select the area near the line $B = \bar{B}$. A depiction of the two pdfs is given in Fig. 5.3. Multiplying the two pdfs combines the information in them. Then we integrate the resulting 2-D pdf to obtain the marginal pdf $\sigma_T(t)$ which represents our best information on T. The MMSE estimate and its variance can then be calculated from $\sigma_T(t)$ (although in general, an estimate and its variance only summarize the information contained in a pdf such as $\sigma_T(t)$). Also

---

[1] Although $\Theta_1(b,t)$ is subject to change on a scale of years or decades, we neglect this phenomenon in this work.

[2] I assume arbitrary constant (i.e., flat) distributions for the remaining distributions in Tarantola's formulation [3, Eq. (1.84)]. Such constant distributions are arguably non-informative (do not add or remove any information from the problem) – see Example 1.15 of [3]. More informative distributions could be selected, but I leave this for future work.

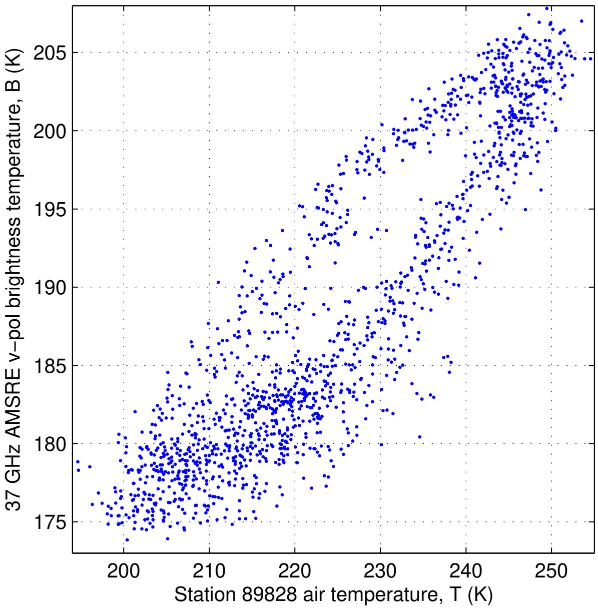

Figure 5.2: All coincident T and B measurements (daily averages) at Antarctic automatic weather station 89828.

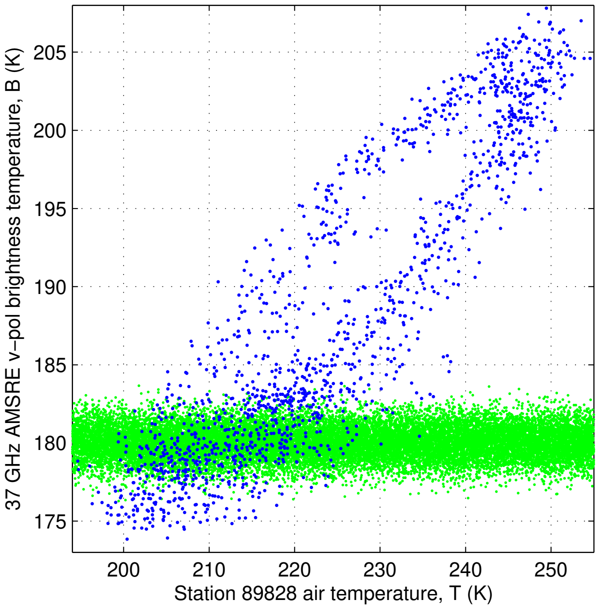

Figure 5.3: Overlay of samples of the two pdfs $\Theta_1(b,t)$ (blue) and $\rho_B(b)$ (green) which are multiplied together and then integrated along the B dimension to obtain the marginal pdf $\sigma_T(t)$ shown in Fig. 5.4.

note that $\sigma_T(t)$ is called the posterior pdf for T and is essentially equivalent to the pdf $p(T|B)$ in a simple Bayesian formulation.

To numerically carry out the operations in (5.3), we start with all available samples of $\Theta_1(b,t)$ (that is, the available coincident T and B data displayed in Fig. 5.2), weight their T values using $\rho_B(b)$, and then add the weighted T values within a bin centered at $T = t$ to find the value of $\sigma_T(t)$.[3] An example of the $\sigma_T(t)$ obtained in this manner is shown in Fig. 5.4, for a measurement value $\bar{B}$ of 180 K with an assumed STD σ of 1 K.

5.3.2 Full Pdf Method: Estimating T from both DOY and B

The sinusoidal method referenced previously employs DOY information by making emissivity a function of DOY. The proposed pdf method utilizes DOY information in the same manner that it uses B data. A pdf $\rho_D(doy)$ is used to convey the DOY on which we desire a T estimate. A Gaussian pdf is used, with STD σ_D and centered at the known DOY on which we wish to estimate T.[4] DOY and B information are used together to estimate T as follows:

$$\sigma_T(t) \;=\; k \int\int \rho_D(doy)\ \rho_B(b)\ \Theta(doy,b,t)\ ddoy\ db. \tag{5.4}$$

Given a particular measurement of B (denoted $\bar{B}$) and the DOY associated with that measurement (denoted $D\bar{O}Y$), (5.4) in more explicit form is

$$\sigma_T(t) \;=\; k_2 \int\int e^{-(doy-D\bar{O}Y)^2/(2\sigma_D^2)}\ e^{-(b-\bar{B})^2/(2\sigma^2)}\ \Theta(doy,b,t)\ ddoy\ db, \tag{5.5}$$

where the scaling coefficients in the Gaussian distributions have been absorbed into the constant k_2. This is the posterior pdf for average air temperature T at the site where $\bar{B}$ was measured on day $D\bar{O}Y$.

---

[3] After repeating this for all desired values of t, the results can be normalized so that $\sigma_T(t)$ integrates to one, as is proper for a pdf. This is equivalent to finding k in (5.3).

[4] If the relationship between DOY and T did not change from year to year, then $\rho_D(doy)$ could have all its mass concentrated at a single point, since the DOY is usually known with complete certainty. But to allow for fluctuation and/or evolution of this relationship, the DOY information is modeled by a non-singular pdf.

This method depends upon two parameters in particular: σ and σ_D. The STD σ_D characterizes variability in air temperature for a fixed DOY as a function of year. For the STD σ, we note that the AMSR-E sensor builders report 0.6 K as the sensitivity of 37 GHz measurements (see the description found at the website http://www.ghcc.msfc.nasa.gov/AMSR/instrument_descrip.html). This can be considered a lower bound on measurement error σ since there are additional error sources for which it does not account (e.g., decorrelation between the surface temperature producing B and air temperature T, azimuth modulation effects [32], and the fact that the AMSRE footprint covers many square kilometers rather than only the point at which the station lies).

The performance of the pdf method depends on the choice of σ and σ_D. The sensitivity of this choice is explored in section 5.4.1. An adaptive method for using training data to choose σ and σ_D is explained in section 5.4.2.

5.4 Application at Three Inland Antarctic Sites

In this section we assess the performance of the pdf method described in the previous section. We compare its estimates with estimates obtained by the sinusoidal method of Shuman et al.

At each of three stations in the interior of Antarctica, three years of quality-controlled concurrent T and B records have been gathered (see Appendix for details). These records are the top three subplots displayed in Fig. 5.5.

In Section 5.4.1, a *single year* of concurrent records from a given site is used to train both the sinusoidal and the pdf method. The methods are then tested on the other two years of concurrent records at the given site. The results are compared at each of the three sites. In Section 5.4.2, *two years* of concurrent records are used to train both methods, while the remaining year is used to test the performance of the two methods at each site.

5.4.1 One Year of Training Data

Sinusoidal Method

To train the sinusoidal method, a brute force numerical search is made to find the sinusoidal emissivity with period 365.25 days which minimizes the RMSE between T estimated from B and the true T, over a period of concurrent T and B data. For example, at station 89828, the minimizing emissivity for the first year of concurrent data is

$$\epsilon = 0.844 + 0.0278 \sin\left((t - 22.3)\frac{2\pi}{365.25}\right), \tag{5.6}$$

where t is in days.

As other examples, the minimizing emissivities for the first year of concurrent data at stations 89813 and 89606 are

$$\epsilon = 0.923 + 0.0077 \sin\left((t - 33.6)\tfrac{2\pi}{365.25}\right) \quad \text{and} \tag{5.7}$$

$$\epsilon = 0.761 + 0.0155 \sin\left((t + 22.2)\tfrac{2\pi}{365.25}\right) \quad , \tag{5.8}$$

respectively.

Such sinusoidal emissivities are then used to estimate T from B for the other years of concurrent data at each station. By using all permutations of training year and test year, six tests are possible at each of the three sites. The RMSE between these estimates and the true T is given in Table 5.1.

Pdf Method

When the pdf method is used for the same training and test data, the estimates depend on the σ and σ_D used. As discussed above, we have a lower bound of $\sigma = 0.6$ K for AMSR-E 37 GHz data but no established value for σ_D. In this subsection, both of these STDs are left as free parameters, with the performance of the pdf method given as a function of them.

To illustrate these dependences, the RMSE of the pdf method is calculated for σ values from 0.6 to 2.4 K in steps of 0.1 K and σ_D values from 1 to 60 days. This

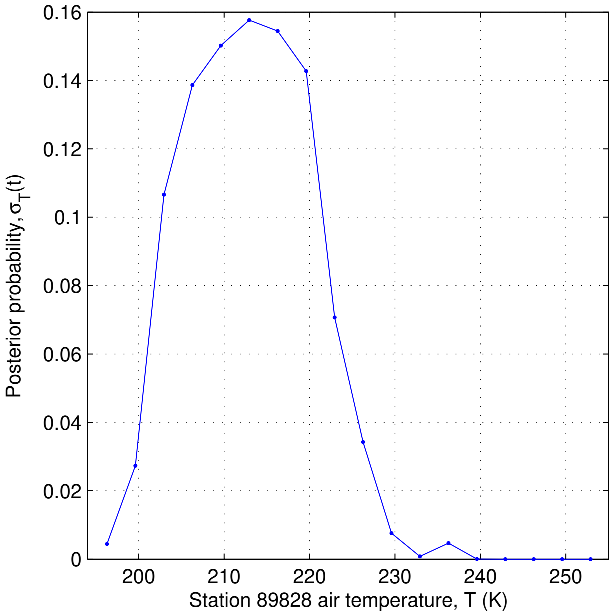

Figure 5.4: Marginal pdf for T obtained from the 2-D pdfs represented in Fig. 5.3.

Table 5.1: RMSE for T using the sinusoidal method at three stations in the interior of Antarctica, with 1 year of training data and 1 year of test data.

	Station 89828	Station 89813	Station 89606
Year 1 = training, year 3 = test	5.7 K	3.6 K	4.3 K
Year 2 = training, year 3 = test	5.5 K	2.5 K	4.3 K
Year 1 = training, year 2 = test	6.4 K	2.2 K	4.3 K
Year 3 = training, year 2 = test	6.3 K	2.4 K	4.4 K
Year 2 = training, year 1 = test	5.0 K	2.1 K	3.9 K
Year 3 = training, year 1 = test	5.1 K	3.4 K	4.1 K
Average at each station:	5.7 K	2.7 K	4.2 K
Average over all stations and tests:		4.2 K	

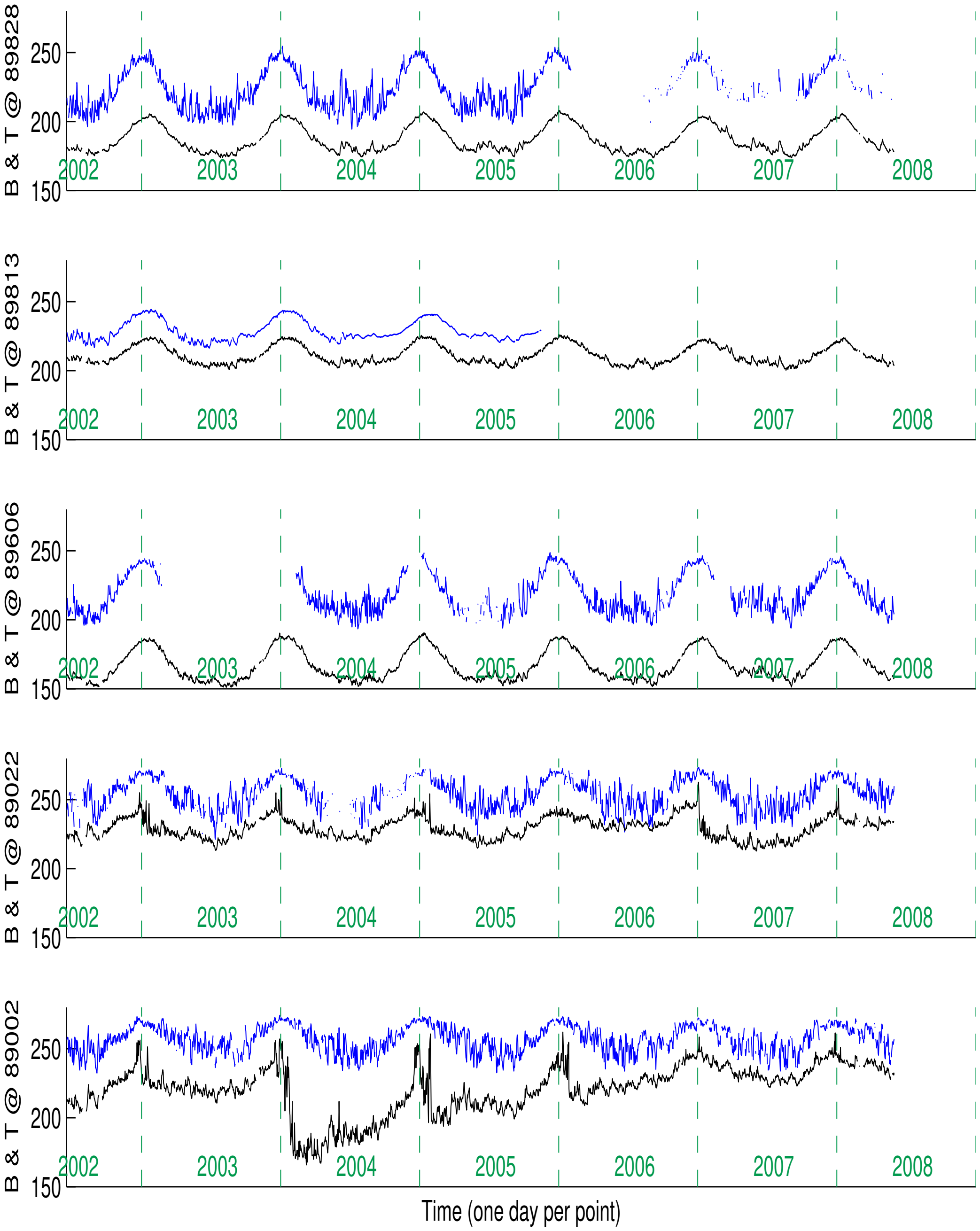

Figure 5.5: (Lower lines, black) brightness temperature B (37 GHz, v-pol) and (higher lines, blue) air temperature T versus time for five Antarctic stations. Both are in units of K.

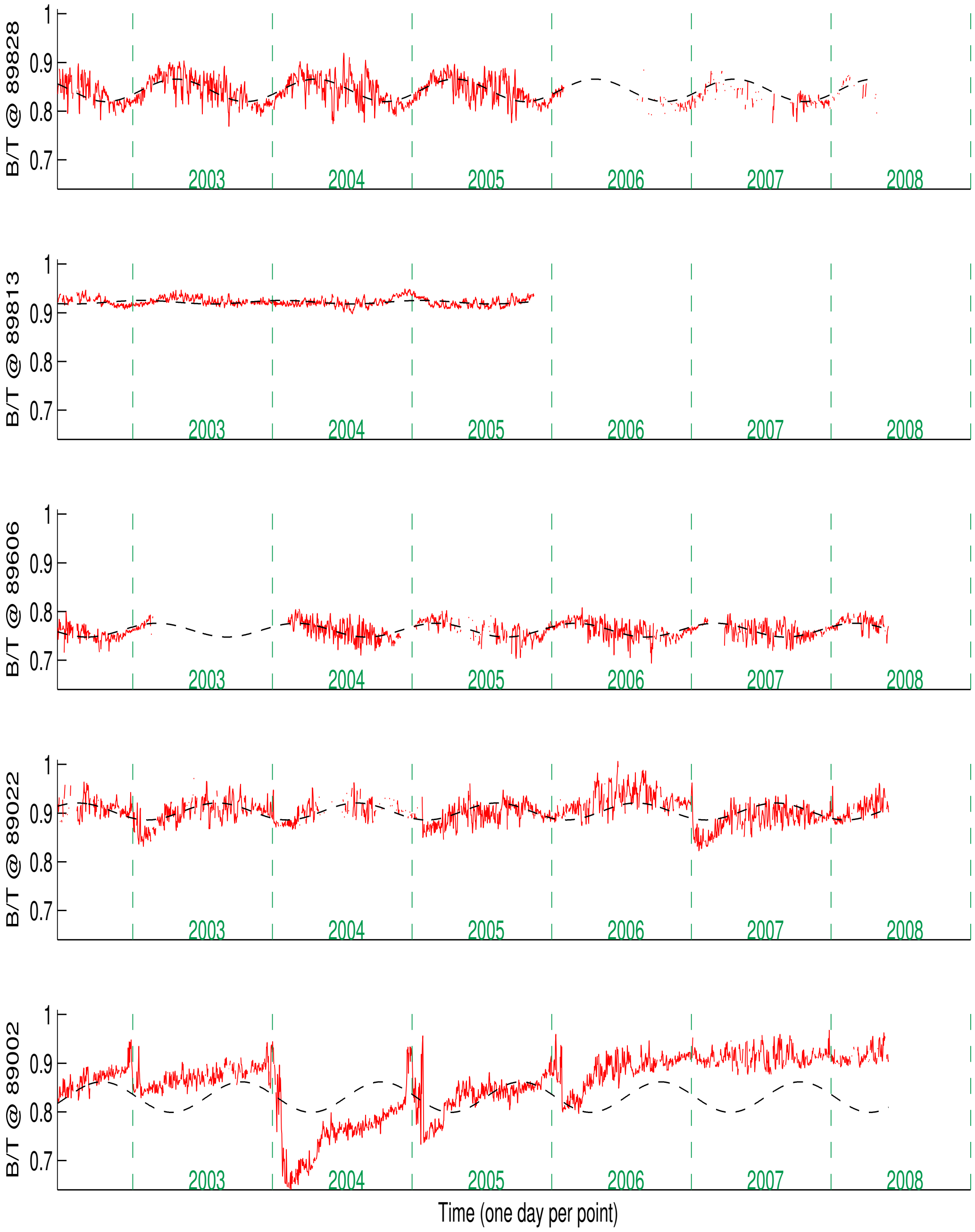

Figure 5.6: Empirical emissivity B/T (red) versus time for the same five Antarctic stations as in Fig. 5.5. Also shown is the sinusoid (black) which best fits the three years used in this work (see Table E.1 in the Appendix for year dates).

is done separately for each of the 18 possible tests. The results at stations 89828, 89813, and 89606 are shown in Fig. 5.7–5.9, respectively. The contours in these plots show the improvement in RMSE which is achieved by the pdf method, in mK relative to the sinusoidal method, as a function of σ and σ_D.

These results indicate that with only one year of training data, the pdf method (with σ and σ_D chosen somewhat blindly) achieves better performance than the sinusoidal method at station 89828 (Dome C). At station 89813, the two methods show similar performance. At station 89606, the two methods also show similar performance. The precise results vary with the particular years used for training and testing.

An average of the improvement over all 18 tests is shown in Fig. 5.10, as a function of σ and σ_D. It reaches a maximum improvement of 0.134 K at $\sigma = 1.4$ K and $\sigma_D = 19$ days. This is an improvement of 3.2% compared with the 4.2 K RMSE which is the average (over the 18 cases) RMSE of the sinusoidal method. Fig. 5.10 also suggests that on average, the pdf method has lower RMSE over a large range of σ and σ_D – roughly over the Cartesian product of $\sigma = 0.6$ to 2.4 K and $\sigma_D = 10$ to 50 days.

5.4.2 Two Years of Training Data

In the following we evaluate and compare the performance of the two methods when two years of training data are available at a site. The sinusoidal method is considered first, followed by the pdf method.

Sinusoidal Method

For the sinusoidal method, the training is the same as with one year of training data except that a minimizing emissivity is found over a two year dataset. The RMSE that results when applying this method to each year of concurrent data at each site are found in Table 5.2 (graphical display of this information is provided in Fig. 5.11). Comparison with the results from only a single year of training (Table

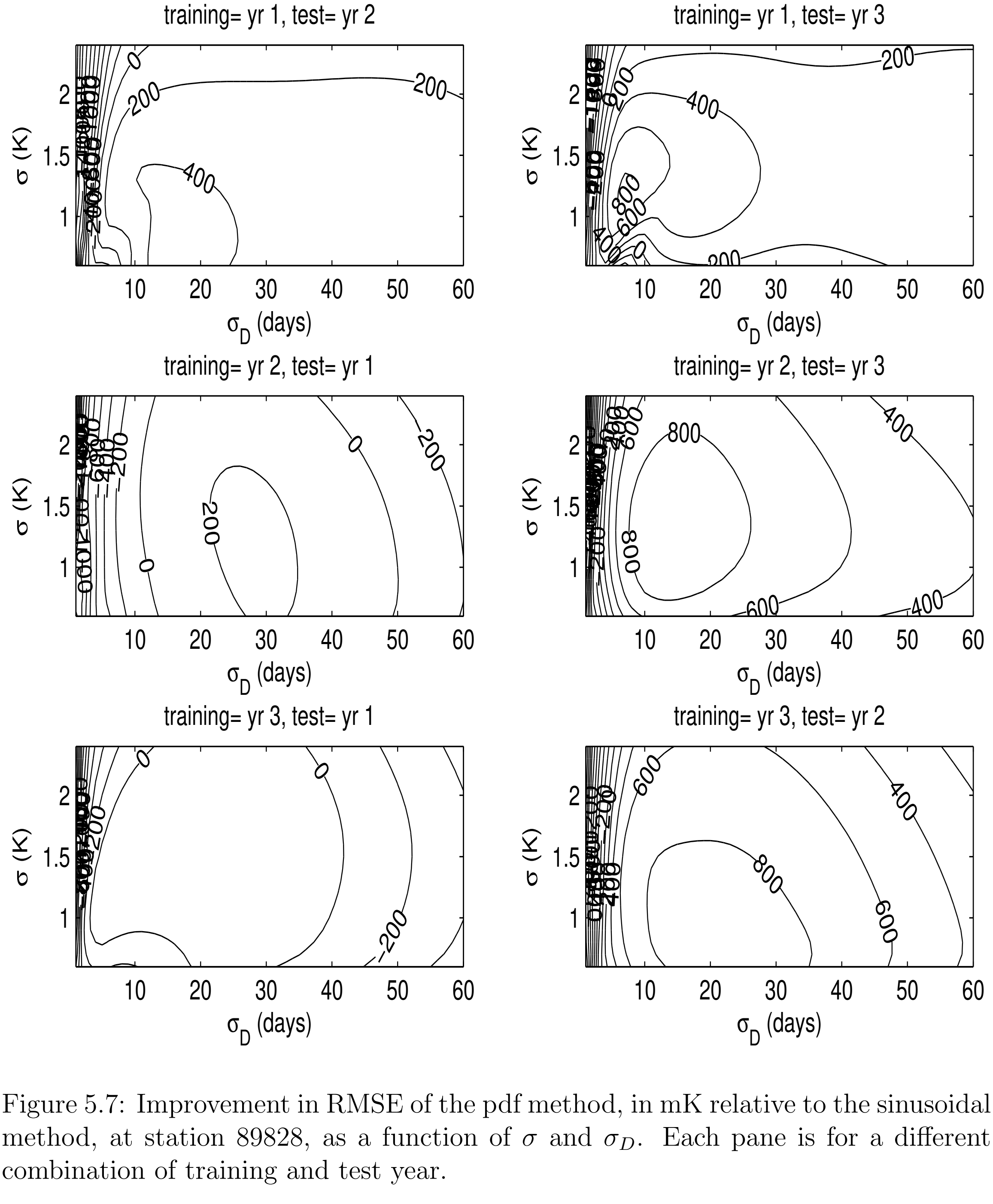

Figure 5.7: Improvement in RMSE of the pdf method, in mK relative to the sinusoidal method, at station 89828, as a function of σ and σ_D. Each pane is for a different combination of training and test year.

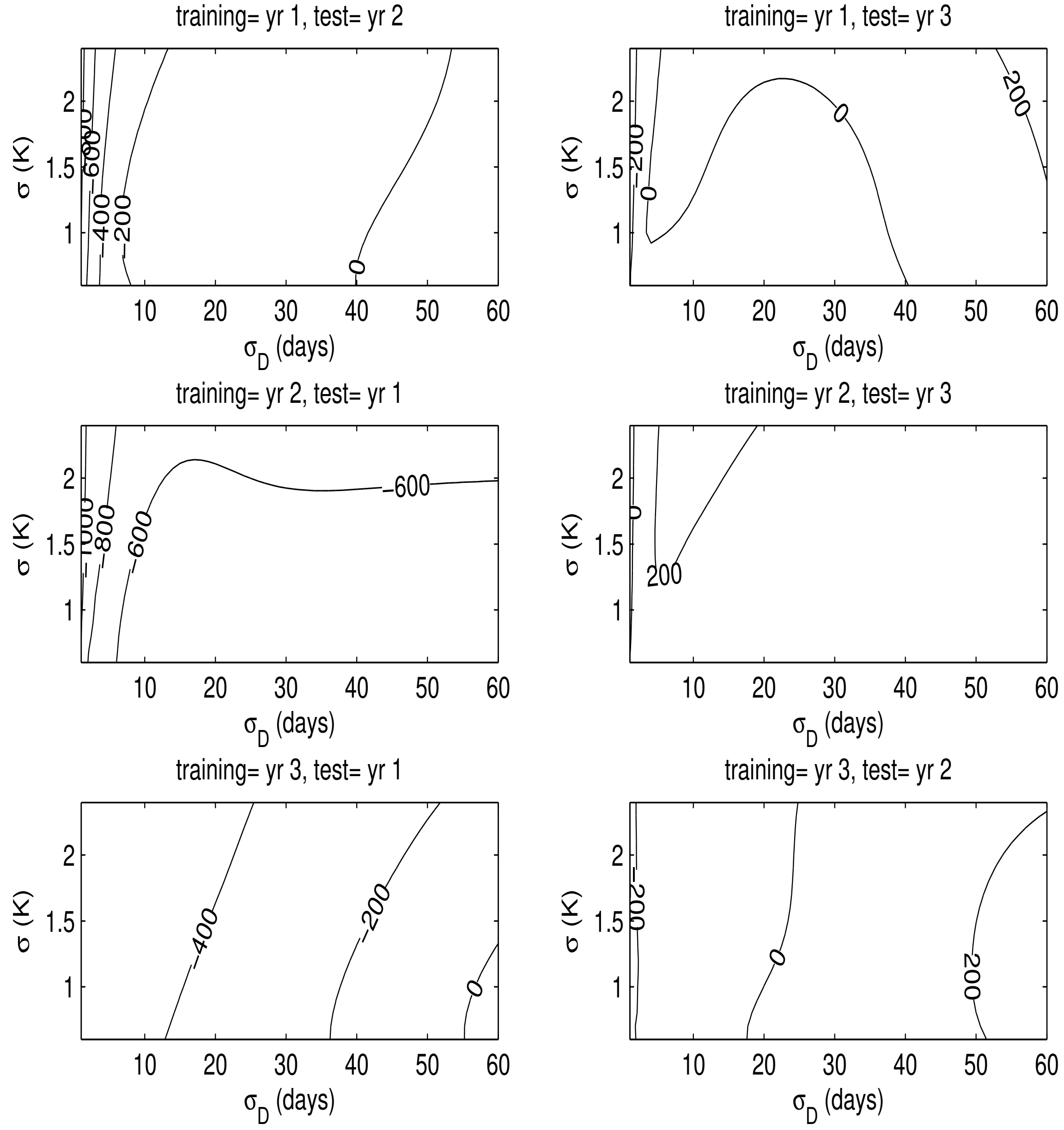

Figure 5.8: Improvement in RMSE of the pdf method, in mK relative to the sinusoidal method, at station 89813, as a function of σ and σ_D. Each pane is for a different combination of training and test year.

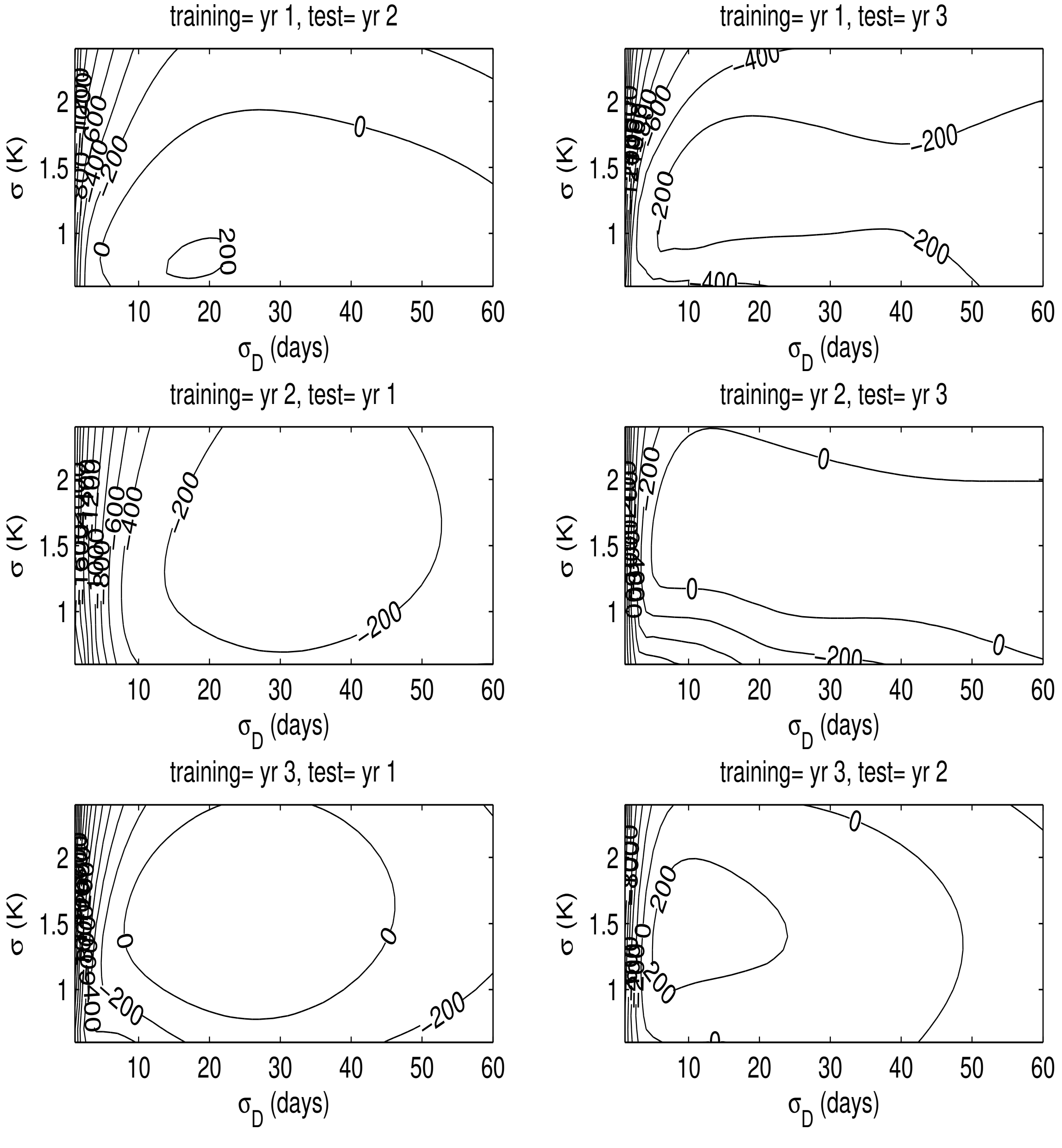

Figure 5.9: Improvement in RMSE of the pdf method, in mK relative to the sinusoidal method, at station 89606, as a function of σ and σ_D. Each pane is for a different combination of training and test year.

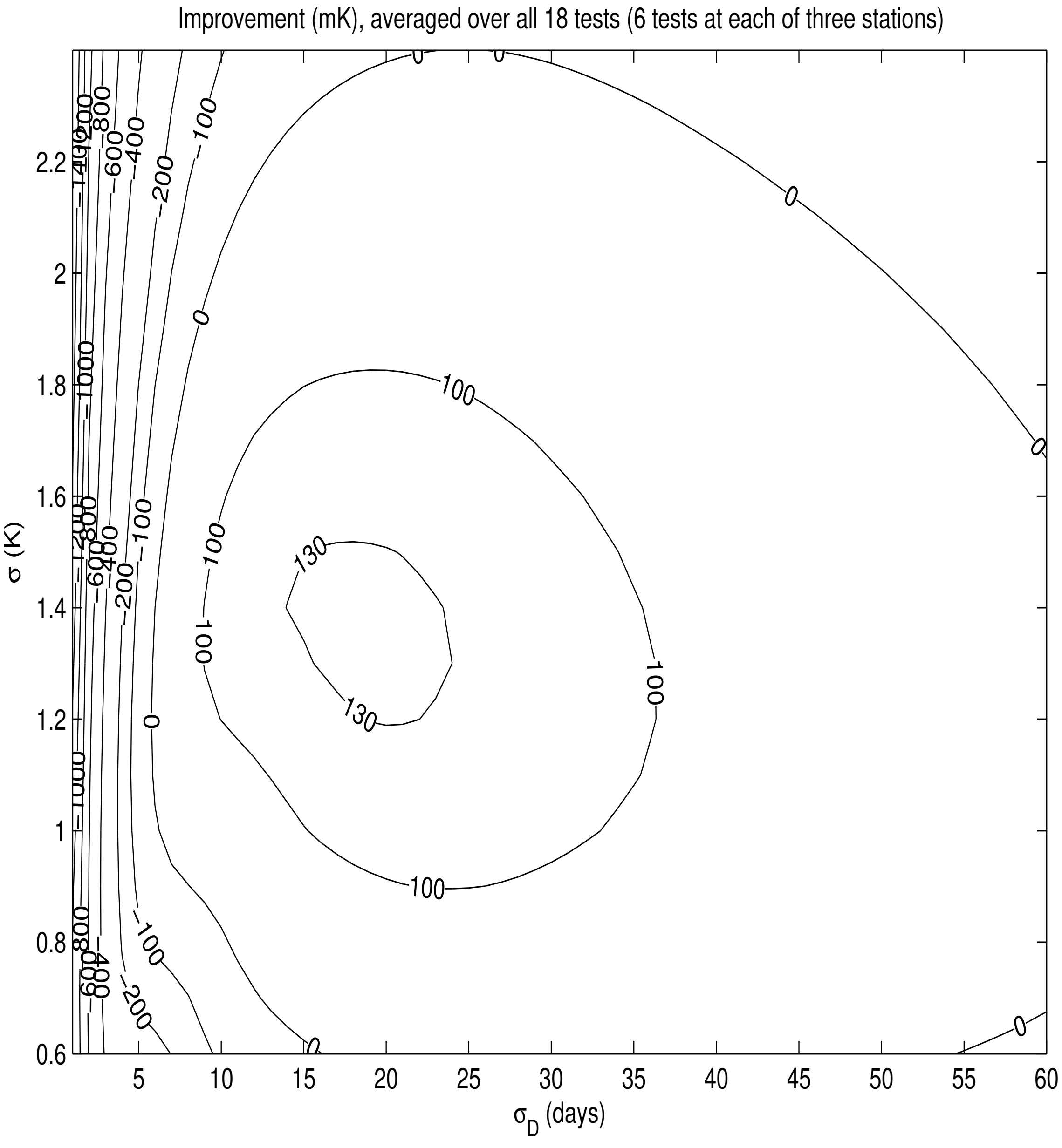

Figure 5.10: Average improvement (over all 18 tests) in RMSE of the pdf method, in mK relative to the sinusoidal method, as a function of σ and σ_D.

Table 5.2: RMSE results of the sinusoidal method and the pdf method with two years of training data and one year of test data at each of three sites in the interior of Antarctica. The σ and σ_D used in the pdf method are chosen by testing the two training years on one another and finding the parameters which maximize performance.

RMSE between Estimates and True (Station) Temperatures

	Station 89828	Station 89813	Station 89606		Station 89022	Station 89002
Years 1 & 2=training, year 3=test:						
Sinusoidal method (K)	5.5	3.0	4.3		7.2	11
Pdf method (K)	4.6	2.6	4.3		5.9	5.6
Improvement (%)	17	15	2		17	48
Years 1 & 3=training, year 2=test:						
Sinusoidal method (K)	6.3	1.7	4.4		12	32
Pdf method (K)	5.6	1.7	4.1		9.5	6.5
Improvement (%)	10	4	6		19	80
Years 2 & 3=training, year 1=test:						
Sinusoidal method (K)	5.1	2.7	4.0		5.9	21
Pdf method (K)	4.8	2.7	4.0		5.7	5.2
Improvement (%)	5	-1	-2		3	76
Ave. improvement (K)	0.61	0.16	0.08		1.2	15.6
Ave. improv't by area	**inland: 0.3 K (7%)**				**coastal: 8.4 K (57%)**	

5.1) indicates that the sinusoidal method does not benefit significantly from a longer training period.

Pdf Method

In contrast to the sinusoidal method, the pdf method *can* benefit significantly from an additional year of training data. The additional year enables an intelligent choice for σ and σ_D as follows.

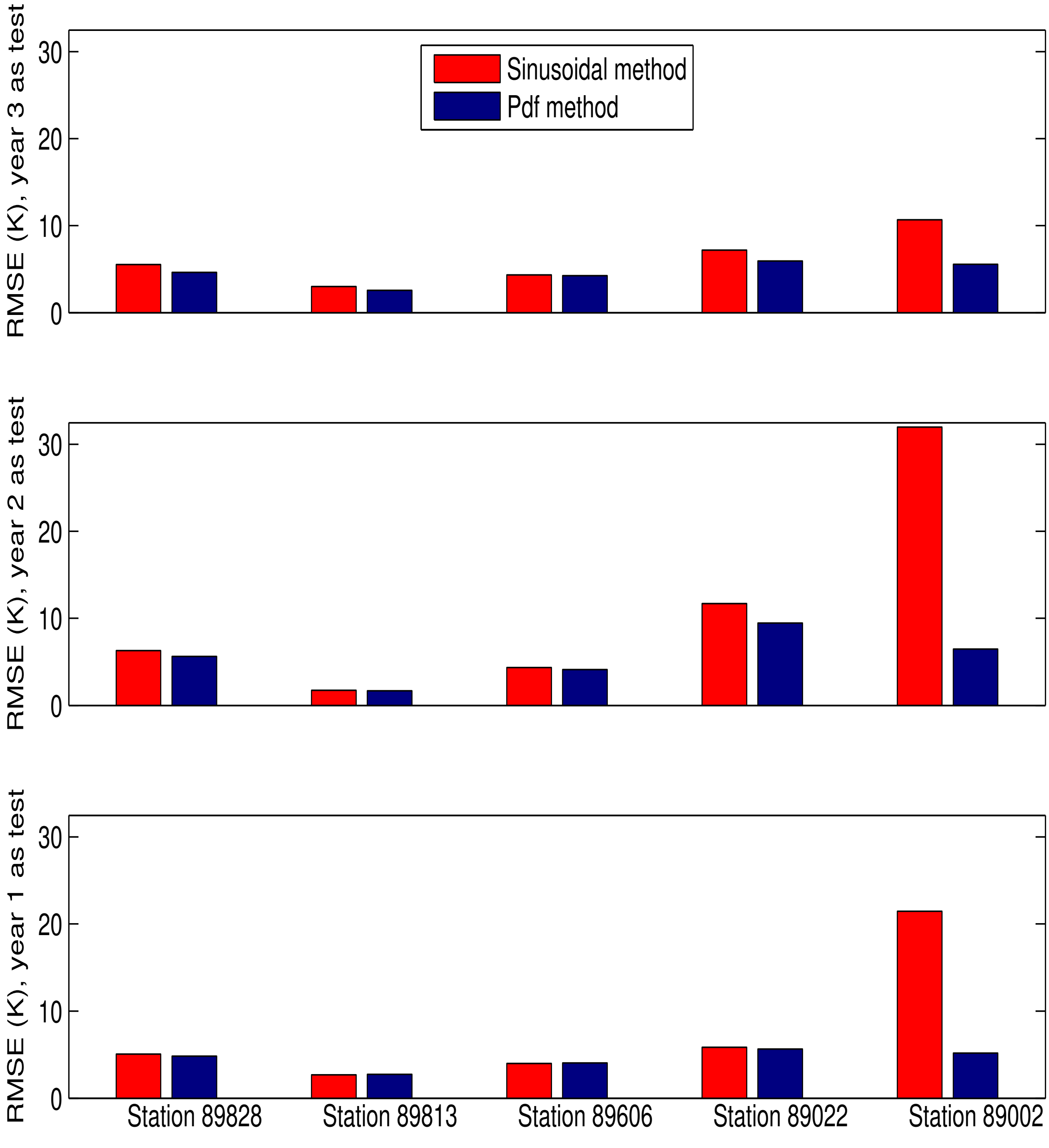

Figure 5.11: Bar graph of same information as in Table 5.2.

For a fixed choice of σ and σ_D, two tests are performed: 1) one using the first year of training data for training and the second year for testing, and 2) the other using the first year of training data for testing and the second year for training. The total RMSE of these two tests is calculated.

This entire process is repeated for many different choices of σ and σ_D, over large ranges of possible values. The choice which minimizes the total RMSE becomes the final choice of σ and σ_D which is used to estimate the third year of data. (Figs. 5.12 – 5.14 include the values selected by this process for each particular site and set of training years.)

The resulting RMSE results are given in the first three columns of Table 5.2 and are graphed in Fig. 5.11. At the inland sites, the pdf method achieves an average RMSE that is 0.3 K (7%) lower than the average RMSE of the sinusoidal method. These results demonstrate that the pdf method is an improvement over the sinusoidal method, at least on the plateau of East Antarctica. Comparisons of true air temperatures with the two estimates are shown in Figs. 5.12–5.14, along with the σ and σ_D values selected by the pdf method.

5.5 Application at Two Coastal Sites

To demonstrate the agility of the pdf method, this section applies the method at two coastal Antarctic sites. T and B data at these sites (Halley, station 89022; and Neumayer, station 89002) are much different from data at inland sites. Plots of T, B, and the associated empirical emissivity at these two stations are given in the lower two panes of Figs. 5.5 and 5.6. The largest difference from inland data is the occurence of melt events, which appear as sharp drops in B near the beginning of most of the years shown.[5]

It is apparent from Fig. 5.6 that a sinusoidal fit is not suited to sites with melt events. Some other fit using a well known function might be contrived, but the

---

[5] Reasons for other more minor differences may be 1) the temperature of the air and of the ice shelves on which the stations lie is more directly influenced by ocean currents and weather than at inland sites, 2) there is greater annual precipitation at coastal sites; and 3) there are differences between the composition of the ice shelves and that of the inland ice sheets – see Appendix E for details on station geography.

pdf method automatically adapts itself optimally. We demonstrate this by applying the sinusoidal and pdf methods as before, with two years of training. The performance attained is given in Table 5.2 and graphed in Fig. 5.11. The estimates themselves are shown in Figs. 5.15-5.16.

From these results, it is apparent that the flexibility of the pdf method can produce large advantages. At station 89022, the pdf method produces better results at almost all times. Its reliance on DOY produces notable robustness in the presence of the melt events (around the beginning of each year). These same melt events cause the sinusoidal method to produce erroneous spikes in estimated T (see Fig. 5.15). Such spikes are especially large in station 89002[6] estimates (Fig. 5.16). Indeed, at station 89002 the sinusoidal method is rendered useless for most of the test times due to 1) the large melt event at the beginning of 2004 and 2) the generally non-sinusoidal nature of the empirical emissivity time series at this station. In contrast, the pdf method estimates show fidelity to smoothed air temperature data.

Additional interesting effects are seen in Fig. 5.16. For Year 1, the pdf method determines that DOY information is more valid than B information, based on the two training years of data. This explains the anomalous rise in pdf method estimates near the middle of 2003 – the rise is shadowing rises in the training years' T for those same DOYs. For Year 2, the pdf method chooses to place greater reliance on B than on DOY, resulting in estimates which show more fluctuation than Year 1 and Year 3 estimates. However, the reliance of the pdf method on DOY is still greater than that of the sinusoidal method – this prevents the pdf method estimates from plunging during and after the melt event, where the sinusoidal method estimates plunge.

5.6 Conclusion

Shuman et al. [29] previously demonstrated a method for estimating Greenland air temperature from 37 GHz v-pol brightness temperature. Their method fits

---

[6] Note that station 89002 is farther north and closer to open ocean than station 89022. This can explain why the melt events are larger at 89002 than at 89022.

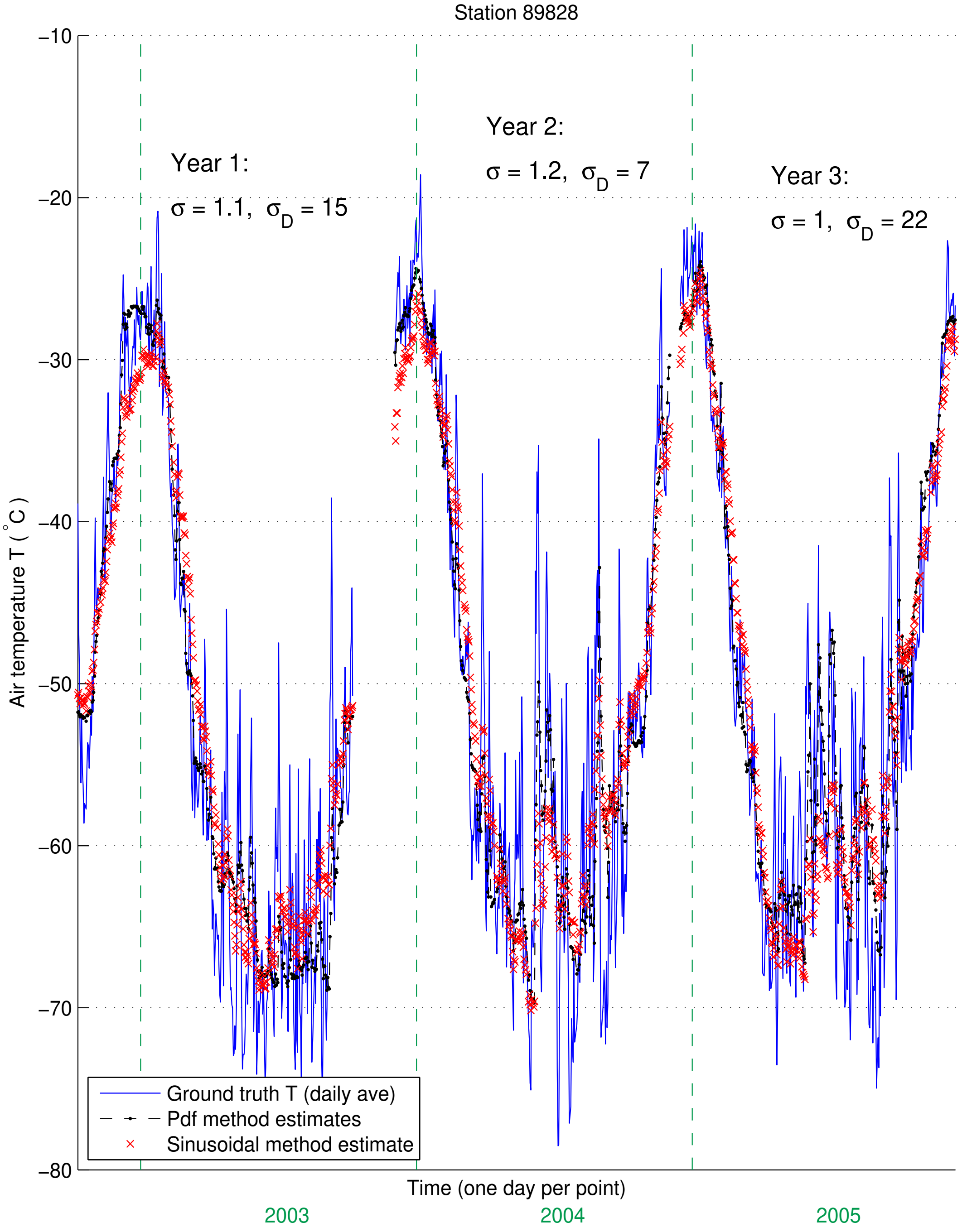

Figure 5.12: True air temperatures (solid blue), estimates of the pdf method (dashed black), and estimates of the sinusoidal method (red x's) at station 89828.

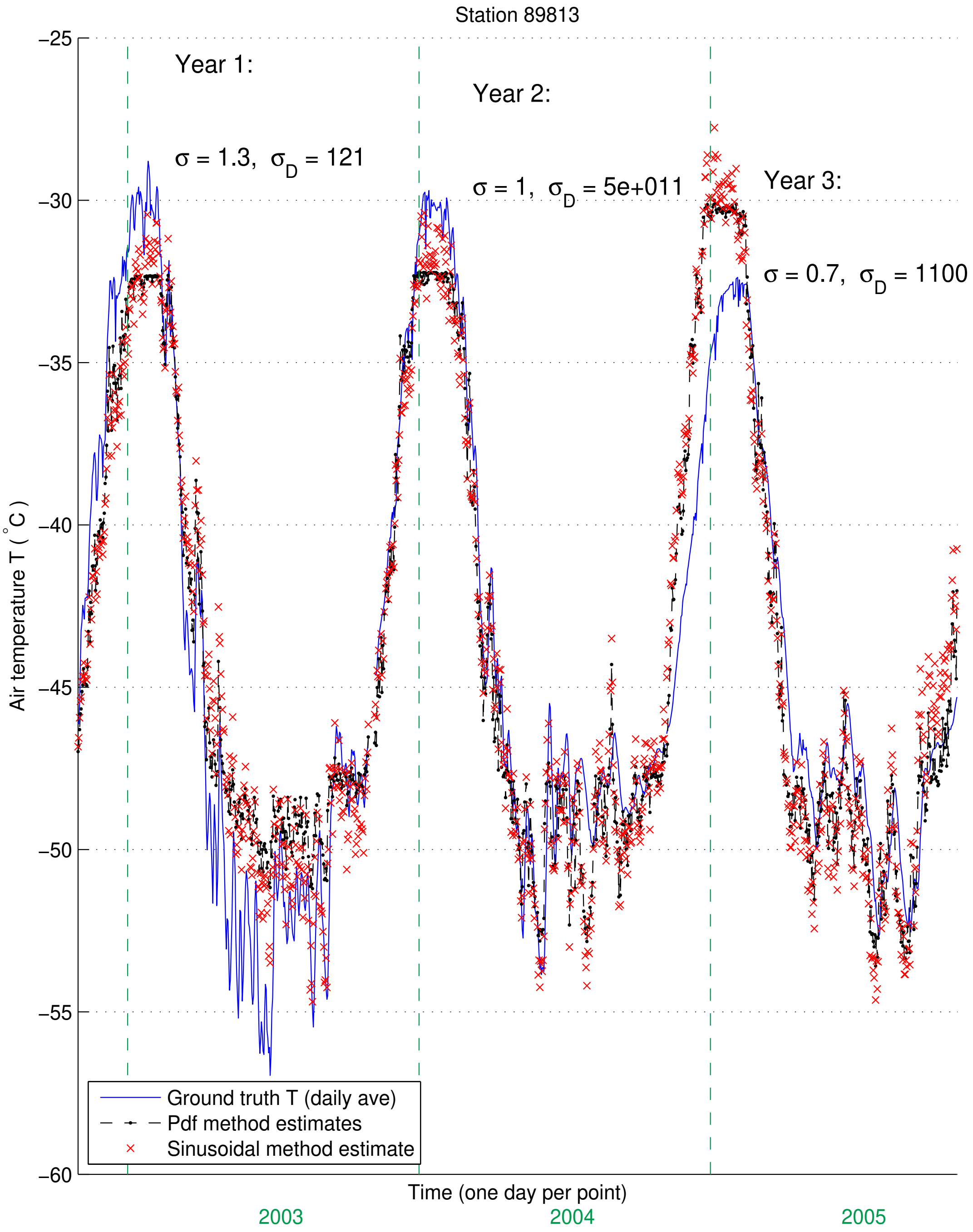

Figure 5.13: True air temperatures (solid blue), estimates of the pdf method (dashed black), and estimates of the sinusoidal method (red x's) at station 89813.

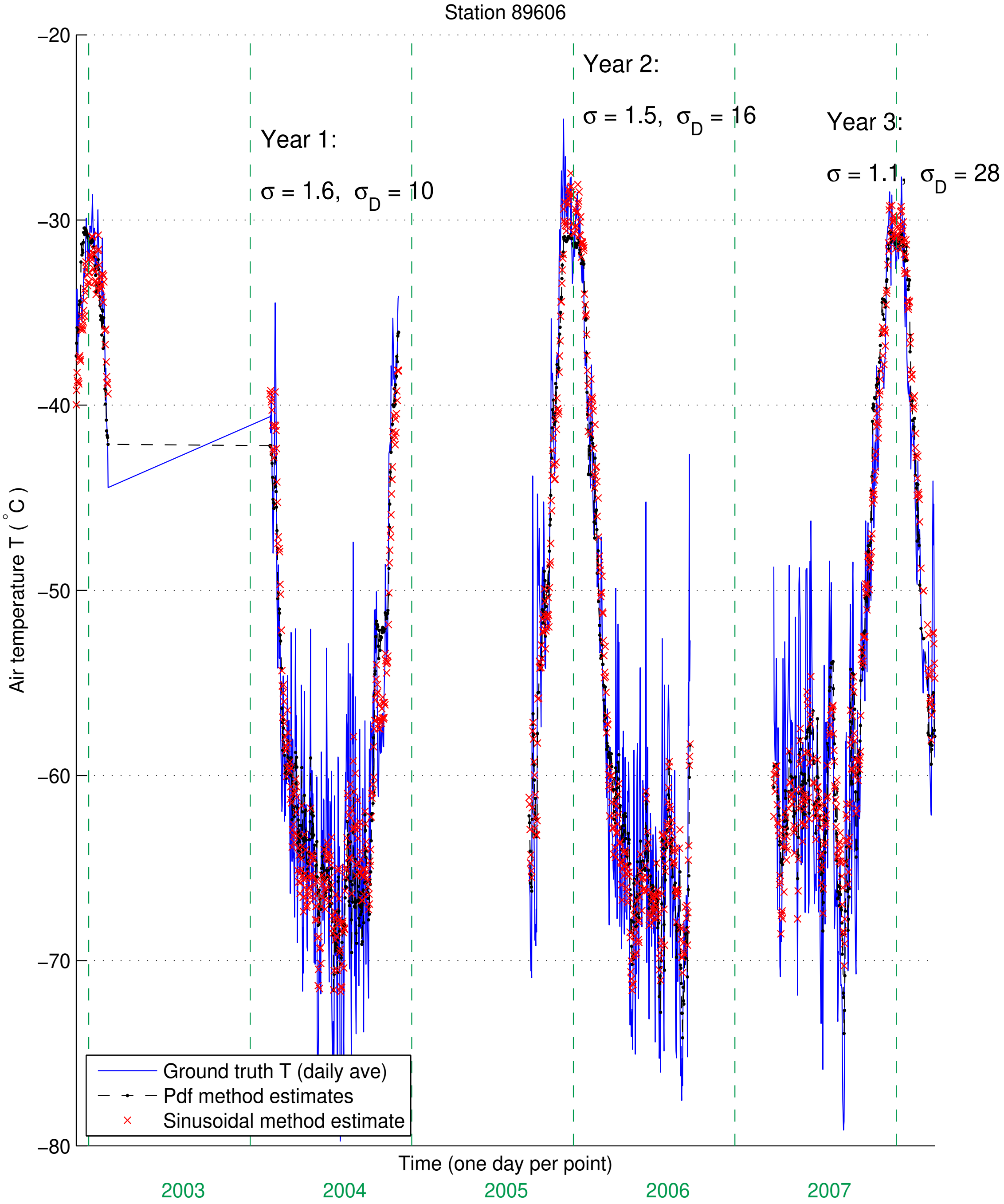

Figure 5.14: True air temperatures (solid blue), estimates of the pdf method (dashed black), and estimates of the sinusoidal method (red x's) at station 89606.

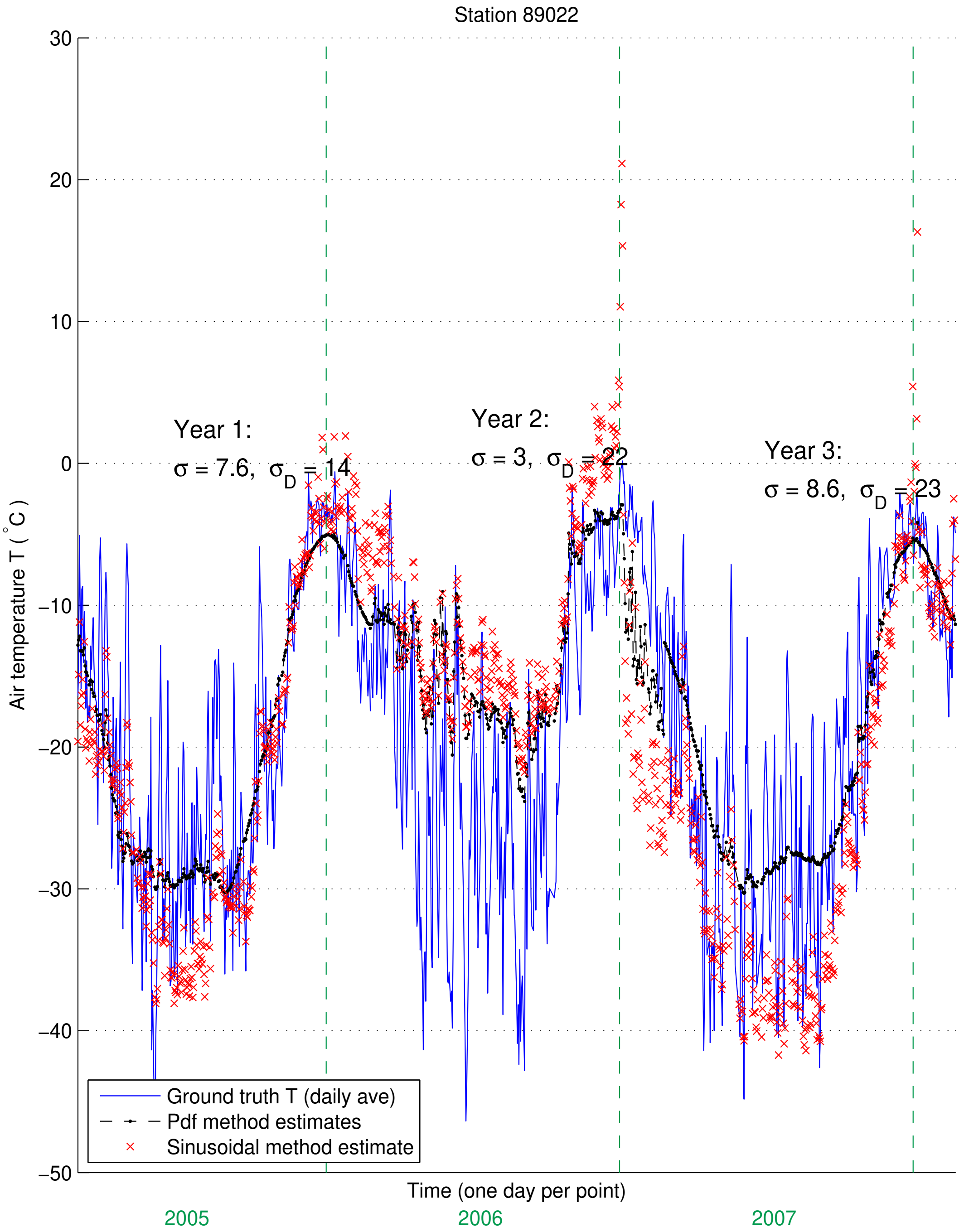

Figure 5.15: True air temperatures (solid blue), estimates of the pdf method (dashed black), and estimates of the sinusoidal method (red x's) at station 89022.

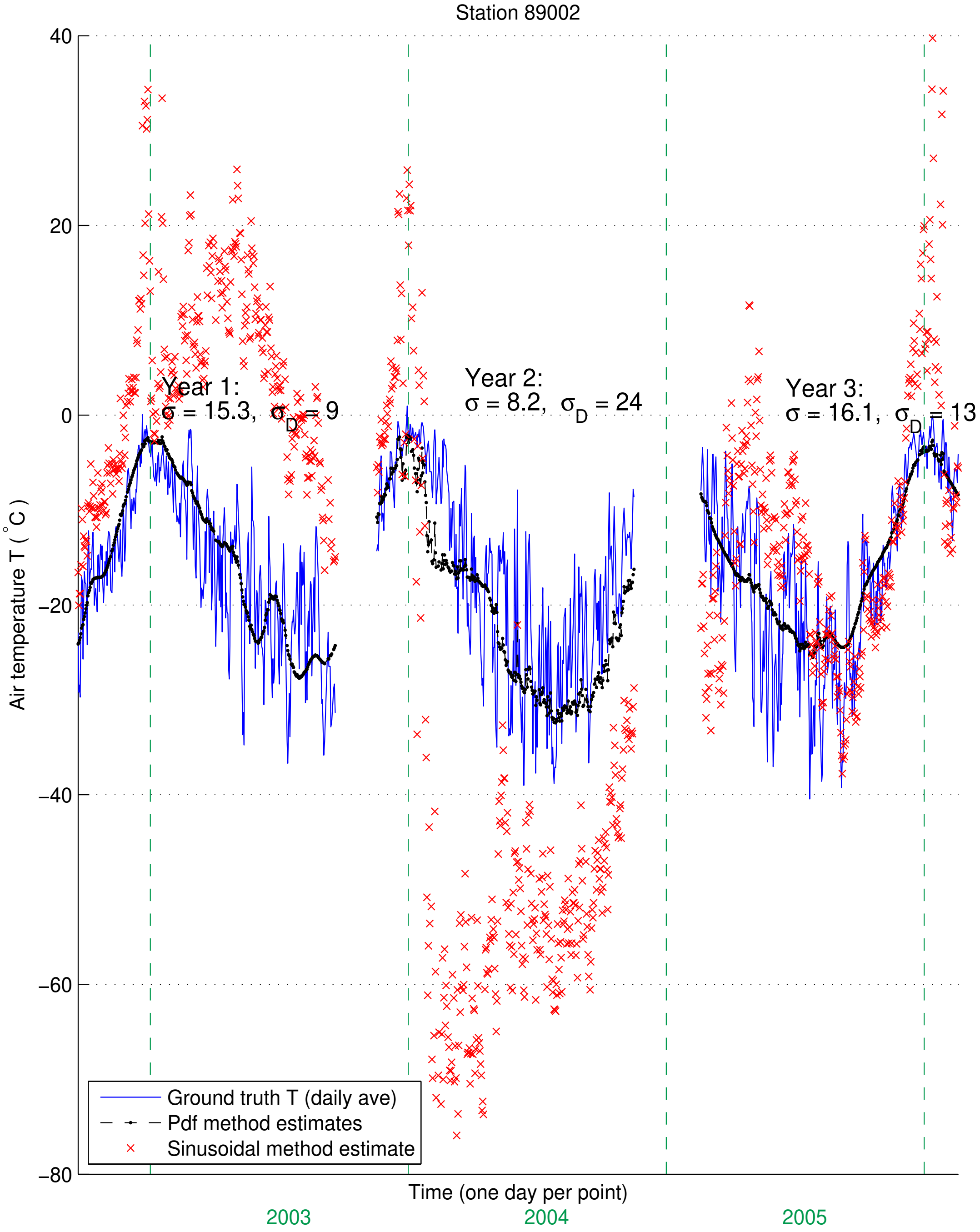

Figure 5.16: True air temperatures (solid blue), estimates of the pdf method (dashed black), and estimates of the sinusoidal method (red x's) at station 89002.

a sinusoid to training data, then uses the sinusoid to estimate air temperature from measurement data. This chapter presents the first demonstrations of this method at sites in Antarctica (five total). The estimates of the sinusoidal method at inland sites have an average RMSE of 4.2 °C when one year of training data is available, decreasing slightly to 4.1 °C when two years of training data are available.

More importantly, this chapter introduces an adaptive alternative, the pdf method. This method uses training data to empirically build a pdf relating measurement data to air temperature. This pdf is then used to translate measurement data to a marginal pdf on air temperature. The mean of the marginal pdf is used as an estimate of air temperature.

With only one year of training data, the performance of the pdf method is slightly better than the sinusoidal method. This holds true over a wide range of two controlling parameters, σ_D and σ. These parameters specify the weight which the pdf method gives to two sources of information (namely, day of year and 37-GHz v-pol brightness temperature measured on that day).

When two years of training data are available, the pdf method can be self-trained to make intelligent choices for σ_D and σ. This intelligence or adaptivity produces significant improvements. Compared to the sinusoidal method, the pdf method reduces RMSE by an average of 0.3 °C at three inland sites and by 8.4 °C at two coastal sites. The large improvement at the coastal sites originates from the adaptability of the pdf method which allows it to anticipate regular melt events.

The pdf method is currently useful for filling in gaps in ground station temperature records and for extending those records to times before or after the station's operating life. Methods for extending the pdf method to areas without ground stations are discussed in the next section.

5.7 Extensions of this Work

Another advantage of the pdf method is that the variance of an air temperature estimate can be estimated even when ground truth data are unavailable. Specifically, the variance can be calculated directly from the marginal pdf for air temperature. A straightforward extension of the present work is to evaluate the quality of this variance estimate by comparing it with RMSE obtained using ground truth. This investigation is left for future work.

In this and previous work, both the pdf and the sinusoidal method require training by ground station truth data. Further study can ascertain the quality of estimates obtained in areas without station data, using data at neighboring sites for training. Currently this prospect is hampered by the lack of additional stations in proximity to the stations whose data is employed in this dissertation.

Another option for estimating temperature in areas without ground stations is to train either the pdf method or sinusoidal method using air temperatures from a new source (such as estimates derived from satellite infrared sensors or numerical weather prediction models). The quality of this approach can be tested by using it at a ground station location and then comparing the resulting estimates with ground station "truth" data. (Note that this option is only meaningful insofar as data from the new source are independent of the ground station data.)

This chapter has focused on minimizing RMSE in air temperature estimates. For applications such as long-term climate change investigation, it may be desirable to minimize bias rather than RMSE. This can readily be accomplished for either the pdf or sinusoidal methods described above. It is done in either case by simply using bias rather than RMSE as the criteria to minimize during the training of the method.

5.8 Acknowledgement

The air temperature data used in this work are provided by the Data Support Section of the Computational and Information Systems Laboratory at the National Center for Atmospheric Research (NCAR). We appreciate the work of Ben Lambert in extracting temperature data from NCAR datasets and in providing Fig. E.1. We also note that NCAR is supported by grants from the National Science Foundation. The original sources of the data may include the Department of Commerce, the

National Oceanic and Atmospheric Administration, the National Weather Service, and the National Center for Environmental Prediction.

Chapter 6

Conclusion

This chapter provides a summary and discussion of the dissertation contributions. A list of publications and suggestions for future work are also provided.

6.1 Principal Contribution

The principal contribution of this dissertation is introducing and proving to the microwave radiometry community that probabilistic approaches to problems can offer significant advantages. This section summarizes these advantages and how the dissertation demonstrates them. Additional contributions are discussed in the next section.

Radiometry is the science of measuring random thermal emissions. Excellent techniques for the characterization and exploitation of the properties of random signals have been developed, some quite recently, in the fields of probability theory and inverse problem theory. This dissertation manifests the value of these techniques in solving three problems of current interest in microwave radiometry.

To begin, in Chapters 3 and 4, probability theory is used to model quantities as random variables rather than simple algebraic unknowns. This permits the inclusion of additional information such as the variances of quantities and their correlations with other quantities. Random variables are more difficult to process than algebraic unknowns. Fortunately, centuries of theoretical work can be tapped to assist in this. In Chapter 3, this approach results in closed-form, analytical expressions for error terms (validated by numerical simulation) whose fidelity is a level higher than previous work. In Chapter 4, this approach results in optimal estimators for calibration parameters. It also leads to the first published information on certain hardware parameters to be obtained from calibration data.

Not only is improvement possible by modeling radiometric signals and error sources as random variables; the outputs of radiometry can also contain more information if they are reported as random variables. It is not uncommon for an estimate to be reported with an error bar or standard deviation. However, covariances between outputs can contain much additional information, as illustrated in Chapter 4. This is particularly relevant as radiometry advances by becoming more polarimetric, in which three or four properties of a radiometric signal are measured simulataneously.

Covariance information can be conveyed by reporting a joint pdf for outputs rather than merely marginal information on each output. In many cases this joint pdf is Gaussian (normal) or nearly so. This is convenient because a Gaussian pdf is easily manipulated. A joint Gaussian pdf is also completely summarized by a vector of means and a covariance matrix, making it computationally inexpensive to store and process. Chapter 4 advocates the greater use of such pdfs as an important next step in the field of microwave radiometry.

Samples of a pdf can serve as either supplements or alternatives to a solution in the form of an analytical pdf. Such samples can be plotted to aid in visualizing available information on parameters, as shown in the figures of Chapter 4. Samples of pdfs can also be used in lieu of analytical pdfs to solve problems where empirical information is used in place of an analytical model, as demonstrated in Chapter 5. Finally, pdfs can be preserved and conveyed by a collection of samples. In terms of subsequent processing, this option can be much more tractable than analytically manipulating non-Gaussian pdfs. While storing and processing many samples of a pdf is computationally more expensive than the conventional choice to simply use the mean, the exponential growth of computing power which has occurred in recent decades puts this option within reach.

6.2 Additional Contributions

As noted above, the principal contribution of this dissertation is introducing and proving to the microwave radiometry community that probabilistic approaches to problems result in significant advantages. The dissertation also illustrates how to work out the difficult details of a probabilistic approach in three diverse situations. These examples can be helpful to others who approach similar situations. This is an important additional contribution.

A number of other, secondary contributions involve specific, current issues or goals of microwave radiometry. These contributions are now described.

6.2.1 Error Analysis of Polarization Rotation Correction

Chapter 3 and Appendix A extend the forward model of polarization rotation to include the random nature of radiation, radiometer channel noises, and (to first order) calibration. With these effects included, derivations are presented for the means, variances, and covariances of radiometer measurements of the first three Stokes parameters (or their modified counterparts) in the presence of polarization rotation. These derivations are validated via Monte Carlo simulation of the original electric-field model.

The error formulas thus derived allow more accurate error analysis and error budgeting than has been possible previously. In particular, they indicate several things about the residual polarization rotation correction (PRC) error. First, the natural third Stokes parameter, of the magnitude expected at L-band for most Earth scenes, is an insignificant source of error compared to $NE\Delta T$.

Second, the dependence of PRC error on rotation angle is determined by residual errors from the calibration process. Since these residuals are unknown (by definition), the dependence of PRC error on rotation angle cannot be predicted as was assumed previously. But if post-launch calibration reduces these residuals to the level of $NE\Delta T$ or less, then the dependence of PRC error on rotation angle is weak – in the limit, the overall PRC error reduces to the $NE\Delta T$ that exists regardless of polarization rotation.

6.2.2 Estimation of Radiometer Calibration and Hardware Parameters

Chapter 4 develops a method of estimating radiometer calibration parameters (channel gains and offsets) which is new to the literature of microwave radiometry. In simulations, it promises to reduce estimation error significantly (by 30% for calibration parameters of the conventional h-pol and v-pol channels; and by factors of 1.5 to 3.7 for calibration parameters of the third Stokes parameter). As a side benefit, the accuracy of the previous method is analyzed and published for the first time.

Chapter 4 and its appendices also derive valuable new information on the estimation of radiometer hardware parameters for the specific class of radiometer under consideration. An explicit formula is derived for obtaining a scattering parameter in terms of calibration measurements. A polarimetric efficiency is shown to be numerically resolvable from calibration measurements. It is proven that for six other hardware parameters (two gains and four detector sensitivies), only ratios or products of the parameters can be resolved. All of these results are valuable for diagnosis of radiometer anomalies.

6.2.3 Improved Polar Air Temperature Estimation

Chapter 5 provides several advances in retrieving polar air temperature from satellite microwave radiometer data. First, the chapter demonstrates a previous empirical method ("sinusoidal method") for the first time on the continent of Antarctica. It shows that the previous method retrieves air temperature at three inland Antarctic sites with a root-mean-square-error (RMSE) of 2.1 to 6.4 K with one year of training data and a RMSE of 1.7 to 6.3 K with two years of training data.

Second, the new, probabilistic empirical method developed in Chapter 5 ("pdf method") provides improved estimation of polar air temperature. At inland Antarctic sites, the reduction in RMSE compared with the previous method is about 0.1 K or 3% with one year of training data and 0.3 K or 7% with two years of training.

An important leap in Chapter 5 is including day-of-year as an important datum to be utilized jointly with radiometer measurements in order to estimate air temperature. This makes the probabilistic method robust at coastal sites. Melt events at coastal sites cause the previous method to be unreliable (RMSE is 5.9 to 32 K, average of 14.8 K, with two years of training) at the two sites with sufficient data to be analyzed in this study. In contrast, the RMSE of the probabilistic method at these sites ranges from 5.2 to 9.5 K (average of 6.4 K) with two years of training data, signifying that the probabilistic method is usable at such sites.

The probabilistic method developed in this chapter can be used immediately for scientific research. First, it can fill in gaps in the temperature records of many polar ground stations. Second, it can extend ground station records to times before or after a station's operating life, provided that satellite data cover those times.

6.3 Publications

Multiple papers publish the work in this dissertation. The material in Chapter 3 appears in the October 2007 issue of the *IEEE Transactions on Geoscience and Remote Sensing* [4]. A summary was presented at a 2006 IEEE conference [33]. Material from Chapter 4 is published in the October 2008 issue of the *IEEE Transactions on Geoscience and Remote Sensing* [5]. The material in Chapter 5 is in final preparation for submission to the peer review and publication process.

6.4 Future Work

A number of extensions can be made to the particular problem solutions worked out in this dissertation. First, the forward model in Chapter 3 can be expanded to include channel gains and uncertainties in them. The work in Chapter 4 provides an excellent start to this extension. Another possible extension of the forward model is the inclusion of antenna pattern nonidealities: sidelobe contributions, which undergo a different amount of polarization rotation than the main beam contribution; and antenna cross-pol contributions, which mix the four Stokes parameters of the scene to some degree.

Next, the findings of Chapter 3 can be validated using data from the Aquarius radiometer when those data become available (the satellite launch is projected to be in 2010). Such validation may then aid in the present design of the SMAP radiometer [34] and/or the development of SMAP calibration algorithms.

Like Chapter 3, the forward model of Chapter 4 is based on a number of stated simplifications (see Section 4.2), and the simulations of Chapter 4 incorporate those simplifications. Future work can assess the effect of these simplifications on the conclusions of Chapter 4 by testing the proposed method on real radiometer data. Alternatively, the forward model can be expanded in an attempt to avoid the simplifications. Also, other profitable extensions can be made by adapting the forward model to other classes of radiometers.

Chapter 4 focuses on estimation of calibration parameters. Future work can assess the improvement in *scene* brightness temperature estimation that can be achieved by using the methodology of Chapter 4 rather than algebraic estimation. This would directly improve science data and is therefore of great interest.

Possible extensions of Chapter 5 are numerous and are discussed in greater detail at the end of that chapter. As a summary, note the following.

First, it would be valuable to explore the accuracy of estimating air temperature in areas without ground stations. One option is to attempt this in areas adjacent to ground stations, using the ground station data for training. Another option is to train either the pdf method or the sinusoidal method using air temperatures from a new source such as estimates derived from satellite infrared sensors or numerical weather prediction models. The quality of this approach can be tested by using it at a ground station location and then comparing the resulting estimates with ground station “truth” air temperature data.

Chapter 5 focuses on minimizing RMSE in air temperature estimates. For applications such as long-term climate change investigation, it may be desirable to minimize bias rather than RMSE. This can readily be accomplished for either the pdf or sinusoidal methods. In both cases, it is done by simply using bias rather than RMSE as the criteria to minimize during the training of the method. A comparison of the resulting biases of the two methods could then be made.

Finally, it is noted in Chapter 5 that the pdf method can predict the variance of its air temperature estimates, without the use of ground truth data. A straightforward extension of the chapter is to evaluate the quality of this variance prediction by comparing it with variance obtained using ground truth.

Appendix A

Derivation of Forward Model of Polarization Rotation

In this Appendix, I derive equations (3.1)-(3.3). These equations comprise the forward model of polarization rotation which is used in Chapter 3. Dr. Jeffrey R. Piepmeier provided the initial sketch of equations (A.1) through (A.15).

A.1 Electric Field Model

Our most basic foundation is a model of the electric fields,

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} \cos\Omega & \sin\Omega \\ -\sin\Omega & \cos\Omega \end{bmatrix} \begin{bmatrix} E_v(t) \\ E_h(t) \end{bmatrix} + \begin{bmatrix} a(t) \\ b(t) \end{bmatrix}. \tag{A.1}$$

$E_v(t)$ and $E_h(t)$ are the components of the total electric field emitted by the scene in the vertical and horizontal directions, respectively (hereafter, our notation suppresses the time dependence, t, of all quantities). Because the number of independent emitters in the scene is large in spaceborne radiometry, E_v and E_h are normally distributed, by the central limit theorem, with zero means [7]. I assume they are real because we are concerned only with the first three Stokes parameter in this work.

E_v and E_h are rotated through an angle Ω, modeling polarization rotation. I consider Ω to be constant over the period of one radiometer measurement. Receiver noise is then added, represented by the electric field amplitudes a and b. Like E_v and E_h, I assume that a and b are normally distributed, zero mean, normal random variables. I also assume they are independent of one another and of E_v and E_h. They represent self emission by the antenna and radiometer. This model neglects sidelobe contributions (as they may undergo different amounts of rotation than the main beam radiation) and cross-coupling of the polarization components caused by the antenna and radiometer non-idealities (cross-pol patterns).

The quantities most commonly reported in radiometry are the first three modified Stokes parameters, as brightness temperatures,

$$\begin{bmatrix} T_v \\ T_h \\ T_U \end{bmatrix} \equiv \begin{bmatrix} < E_v^2 > \\ < E_h^2 > \\ 2 < E_v E_h > \end{bmatrix}, \tag{A.2}$$

to which I add, for this document,

$$\begin{bmatrix} T_{RX,v} \\ T_{RX,h} \end{bmatrix} \equiv \begin{bmatrix} < a^2 > \\ < b^2 > \end{bmatrix}. \tag{A.3}$$

In these and subsequent definitions, I ignore a proportionality constant which converts the product of two electric fields to a brightness temperature.[1]

A quantity of high interest to users of radiometry data is the second Stokes parameter, $T_Q \equiv T_v - T_h \equiv < E_v^2 > - < E_h^2 >$ where $< \cdot >$ denotes the expected value (ensemble average). In addition to the definition in (A.2), T_U can be equivalently defined in a manner analogous to the definition of T_Q. This definition is $T_U \equiv T_{+45} - T_{-45}$, where T_{+45} is the brightness temperature of the component of the incident radiation that is linearly polarized at 45° with respect to the E_v and E_h axes.

Our model assumes a radiometer architecture in which the signals at +45° and −45° linear polarization (in the radiometer polarization basis) are synthesized from x and y after enough amplification of x and y (by LNAs) that receiver noise added after this synthesis is negligible. Radiometers which create the signals at +45° and −45° earlier (such as from direct measurement of T_{+45} and T_{-45}) require that additional noise terms be added to the additional channels. This would add many terms to the final forward model and the error formulas.

---

[1] This conversion also assumes a narrow band radiometer, so that the frequency spectrums of $E_v(t)$, $E_h(t)$, $a(t)$, and $b(t)$ are flat, see [6].

A.2 Description of Parameters

In this document, I could express our results in terms of $T_v, T_h, T_{RX,v}$, and $T_{RX,h}$. It is more concise, however, to use the related quantities $T_I \equiv T_v + T_h$, $T_Q \equiv T_v - T_h$, $T_{RX,I} \equiv T_{RX,v} + T_{RX,h}$, and $T_{RX,Q} \equiv T_{RX,v} - T_{RX,h}$. Note that T_I, T_Q, and T_U comprise the first three Stokes parameters [15] as brightness temperatures. I also note that in the final expressions for bias and variance (and hence RMSE), T_I and $T_{RX,I}$ always appear added together, never separately. Therefore I reduce our parameter set by using $T_{sys,I} \equiv T_I + T_{RX,I}$.

Beside $T_I, T_Q, T_U, T_{RX,I}$, and $T_{RX,Q}$, other parameters are Ω, N, $\Delta T_{RX,I}$, $\Delta T_{RX,Q}$, and $\Delta T_{RX,U}$ (N is defined early in Section A.3; $\Delta T_{RX,I}$, $\Delta T_{RX,Q}$, and $\Delta T_{RX,U}$ are defined in Section A.5). This collection of ten parameters can be used to completely describe the forward problem and I therefore refer to them as the "original parameters." Other quantities are defined for convenience but can be expressed in terms of these original ten.

The symbols x and y represent the electric fields to be detected by the radiometer. By the construction of (A.1), they are also zero mean, normal random variables. I denote their expected squared values, as brightness temperatures, with

$$\begin{bmatrix} T_{sys,v} \\ T_{sys,h} \\ T_{sys,U} \end{bmatrix} \equiv \begin{bmatrix} < x^2 > \\ < y^2 > \\ 2 < xy > \end{bmatrix}. \tag{A.4}$$

Using (A.2), (A.3), and the facts that a and b are independent of all other quantities and are zero mean, we find

$$\begin{aligned} T_{sys,v} &= \left\langle (E_v \cos\Omega + E_h \sin\Omega + a)^2 \right\rangle \\ &= T_v \cos^2\Omega + T_h \sin^2\Omega + \frac{T_U}{2}\sin 2\Omega + T_{RX,v}. \end{aligned} \tag{A.5}$$

By a similar process,

$$\begin{aligned} T_{sys,h} &= T_h \cos^2\Omega + T_v \sin^2\Omega - \frac{T_U}{2}\sin 2\Omega + T_{RX,h}, \\ && \text{(A.6)} \\ T_{sys,U} &= -T_Q \sin 2\Omega + T_U \cos 2\Omega. && \text{(A.7)} \end{aligned}$$

A.3 Measured Temperatures, $\hat{T}_{sys,v}$, $\hat{T}_{sys,h}$ and $\hat{T}_{sys,U}$

A conventional two-channel radiometer measures $T_{sys,v}$ and $T_{sys,h}$ by a time average,

$$\hat{T}_{sys,v} \equiv \frac{1}{\tau}\int_0^\tau x^2 dt,\ \hat{T}_{sys,h} \equiv \frac{1}{\tau}\int_0^\tau y^2 dt. \tag{A.8}$$

I use hats to denote measured or estimated quantities, which are random variables, as opposed to the unhatted quantities which represent the desired true quantities, such as the ensemble average of a random variable.

A three-channel polarimetric radiometer also measures

$$\hat{T}_{sys,U} \equiv \frac{2}{\tau}\int_0^\tau xy\ dt. \tag{A.9}$$

As shown in [7], $\hat{T}_{sys,v}$, $\hat{T}_{sys,h}$, and $\hat{T}_{sys,U}$ can be rewritten as sums of independent samples,

$$\hat{T}_{sys,v} = \frac{1}{N}\sum_{i=1}^{N} x_i^2,\ \hat{T}_{sys,h} = \frac{1}{N}\sum_{i=1}^{N} y_i^2,\ \hat{T}_{sys,U} = \frac{2}{N}\sum_{i=1}^{N} x_i y_i, \tag{A.10}$$

where $N = 2B\tau$, B is the sensor bandwidth, and τ is the integration time.

I next proceed to find the distributions of $\hat{T}_{sys,v}$, $\hat{T}_{sys,h}$ and $\hat{T}_{sys,U}$. For large N (for Aquarius, $N \approx 480,000,000$), $\hat{T}_{sys,v}$ is so nearly Gaussian, by the central limit theorem, that I assume it is Gaussian. Similar results apply for $\hat{T}_{sys,h}$ and $\hat{T}_{sys,U}$. Therefore, they can be very well characterized by only their means, variances, and covariances, which I derive next.

A.3.1 Means of $\hat{T}_{sys,v}$, $\hat{T}_{sys,h}$ and $\hat{T}_{sys,U}$

The ensemble average (expected value) of $\hat{T}_{sys,v}$ is

$$< \hat{T}_{sys,v} >= \left\langle \frac{1}{N}\sum_{i=1}^{N} x_i^2 \right\rangle = \frac{1}{N}\sum_{i=1}^{N} < x_i^2 >= T_{sys,v}. \tag{A.11}$$

Similarly, $< \hat{T}_{sys,h} >= T_{sys,h}$ and $< \hat{T}_{sys,U} >= T_{sys,U}$.

A.3.2 $Var(\hat{T}_{sys,v})$ and $Var(\hat{T}_{sys,h})$

$$\begin{aligned} Var(\hat{T}_{sys,v}) &= \left\langle (\frac{1}{N}\sum_{i=1}^{N} x_i^2)(\frac{1}{N}\sum_{j=1}^{N} x_j^2) \right\rangle - T_{sys,v}^2 \\ &= \left\langle \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} x_i^2 x_j^2 \right\rangle - T_{sys,v}^2, \end{aligned} \tag{A.12}$$

which we separate into terms for which $i \neq j$ and for which $i = j$:

$$= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1(\neq i)}^{N} < x_i^2 x_j^2 > + \frac{1}{N^2}\sum_{i=1}^{N} < x_i^4 > - T_{sys,v}^2. \tag{A.13}$$

Using the independence of samples i and j and the known fourth moment of zero-mean normal random variables,

$$\begin{aligned} &= \frac{1}{N^2}\sum_{i=1}^{N} < x_i^2 > \sum_{j=1(\neq i)}^{N} < x_j^2 > \\ &+ \frac{1}{N^2}\sum_{i=1}^{N} 3 < x_i^2 >^2 - T_{sys,v}^2. \end{aligned} \tag{A.14}$$

Then, using (A.4),

$$Var(\hat{T}_{sys,v}) = \frac{2}{N} T_{sys,v}^2. \tag{A.15}$$

By a similar process,

$$Var(\hat{T}_{sys,h}) = \frac{2}{N} T_{sys,h}^2. \tag{A.16}$$

A.3.3 $Var(\hat{T}_{sys,U})$

By a process similar to (A.12) through (A.15)

$$Var(\hat{T}_{sys,U}) = \frac{-T_{sys,U}^2}{N} + \frac{4}{N} < x^2 y^2 >, \tag{A.17}$$

Consider $< x^2 y^2 >$ alone. Using the definitions of x and y in (A.1), it can be expanded to several dozen terms. The independence of a and b from E_v and E_h means that many terms can be factored as $< a >$, $< b^2 >$, and so on. Then using (A.2), (A.3), the fact that a and b are zero mean, and the known fourth moment of zero-mean normal random variables, many terms drop out or simplify, leaving

$$\begin{aligned} < x^2y^2 > &= \frac{1}{2} < E_vE_h^3 - E_v^3E_h > \sin 4\Omega \\ &+ \left(\frac{3}{4} < E_v^2E_h^2 > -\frac{3}{8}(T_v^2 + T_h^2)\right) \cos 4\Omega \\ &- \frac{1}{2}T_{RX,Q}(T_U \sin 2\Omega + T_Q \cos 2\Omega) \\ +\frac{1}{4} < E_v^2E_h^2 > &+ \frac{3}{8}(T_v^2 + T_h^2) + T_{RX,v}T_{RX,h} + \frac{1}{2}T_IT_{RX,I}. \end{aligned} \tag{A.18}$$

E_v and E_h are marginally zero-mean Gaussians, with variances of T_v and T_h and a covariance of $T_U/2$. Assuming they are jointly Gaussian, their joint probability density function (pdf) is completely specified. We can therefore determine $< E_v^3E_h >$, $< E_vE_h^3 >$, and $< E_v^2E_h^2 >$ by direct integration:

$$\begin{aligned} < E_v^3E_h > &= \frac{1}{\pi\sqrt{4T_vT_h - T_U^2}} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E_v^3E_h e^{\frac{-2T_hE_v^2 + 2T_UE_hE_v - 2T_vE_h^2}{4T_vT_h - T_U^2}} & \, dE_v \, dE_h. \end{aligned} \tag{A.19}$$

Using a table of integrals [35], the known second and fourth moments of zero-mean Gaussians, and much algebra, this reduces to

$$< E_v^3E_h > = \frac{3}{2}T_UT_v. \tag{A.20}$$

By similar processes, we find

$$< E_vE_h^3 > = \frac{3}{2}T_UT_h \text{ and } < E_v^2E_h^2 > = T_vT_h + \frac{1}{2}T_U^2. \tag{A.21}$$

By using these results in (A.18) and then using (A.18) in (A.17), we obtain, after much algebraic manipulation,

$$Var(\hat{T}_{sys,U}) = \frac{1}{N}[T_{sys,I}^2 - T_{sys,Q}^2 + T_{sys,U}^2], \tag{A.22}$$

where $T_{sys,I} \equiv T_{sys,v} + T_{sys,h} = T_I + T_{RX,I}$ and $T_{sys,Q} \equiv T_{sys,v} - T_{sys,h}$.

A.3.4 Covariances of $\hat{T}_{sys,v}$, $\hat{T}_{sys,h}$, and $\hat{T}_{sys,U}$

We wish to determine the covariances that exist between $\hat{T}_{sys,v}$, $\hat{T}_{sys,h}$, and $\hat{T}_{sys,U}$. Similar to the derivation of (A.22), it can be shown that

$$Cov(\hat{T}_{sys,v}, \hat{T}_{sys,h}) = \frac{T_{sys,U}^2}{2N}, \tag{A.23}$$

$$Cov(\hat{T}_{sys,v}, \hat{T}_{sys,U}) = \frac{2T_{sys,v}T_{sys,U}}{N}, \tag{A.24}$$

$$Cov(\hat{T}_{sys,h}, \hat{T}_{sys,U}) = \frac{2T_{sys,h}T_{sys,U}}{N}. \tag{A.25}$$

A.4 Definition and Characterization of $\hat{T}_{sys,I}$ and $\hat{T}_{sys,Q}$

It is more convenient to work with the sum and difference of $\hat{T}_{sys,v}$ and $\hat{T}_{sys,h}$ than with these quantities themselves. Therefore we define $\hat{T}_{sys,I} \equiv \hat{T}_{sys,v} + \hat{T}_{sys,h}$ and $\hat{T}_{sys,Q} \equiv \hat{T}_{sys,v} - \hat{T}_{sys,h}$. Using the formulas given above, it is straightforward to show that

$$< \hat{T}_{sys,I} >= T_I + T_{RX,I} \equiv T_{sys,I}, \tag{A.26}$$

$$< \hat{T}_{sys,Q} >= T_Q \cos 2\Omega + T_U \sin 2\Omega + T_{RX,Q} \equiv T_{sys,Q}, \tag{A.27}$$

and that the variances and covariances of $\hat{T}_{sys,I}$, $\hat{T}_{sys,Q}$, and $\hat{T}_{sys,Q}$ can be summarized with the symmetric covariance matrix

$$N \cdot \begin{bmatrix} Var(\hat{T}_{sys,I}) & Cov(\hat{T}_{sys,I}, \hat{T}_{sys,Q}) & Cov(\hat{T}_{sys,I}, \hat{T}_{sys,U}) \\ & Var(\hat{T}_{sys,Q}) & Cov(\hat{T}_{sys,Q}, \hat{T}_{sys,U}) \\ & & Var(\hat{T}_{sys,U}) \end{bmatrix} =$$

$$\begin{bmatrix} T_{sys,I}^2 + T_{sys,Q}^2 + T_{sys,U}^2 & 2T_{sys,I}T_{sys,Q} & 2T_{sys,I}T_{sys,U} \\ & T_{sys,I}^2 + T_{sys,Q}^2 - T_{sys,U}^2 & 2T_{sys,Q}T_{sys,U} \\ & & T_{sys,I}^2 - T_{sys,Q}^2 + T_{sys,U}^2 \end{bmatrix}. \tag{A.28}$$

A.5 Forward Model of Rotated and Calibrated Brightness Temperatures

As discussed at the beginning of Section A.3, the measured temperatures are normal random variables with the means and variances just found. It is convenient to break them up into the sum of their means and zero-mean, normal random variables,

$$\begin{bmatrix} \hat{T}_{sys,v} \\ \hat{T}_{sys,h} \\ \hat{T}_{sys,U} \end{bmatrix} \equiv \begin{bmatrix} T_{sys,v} + \Delta T_{sys,v} \\ T_{sys,h} + \Delta T_{sys,h} \\ T_{sys,U} + \Delta T_{sys,U} \end{bmatrix}, \tag{A.29}$$

and similarly for the quantities defined for convenience,

$$\hat{T}_{sys,I} \equiv T_{sys,I} + \Delta T_{sys,I}, \tag{A.30}$$

$$\hat{T}_{sys,Q} \equiv T_{sys,Q} + \Delta T_{sys,Q}, \tag{A.31}$$

where $\Delta T_{sys,I} \equiv \Delta T_{sys,v} + \Delta T_{sys,h}$ and $\Delta T_{sys,Q} \equiv \Delta T_{sys,v} - \Delta T_{sys,h}$.

Expanding these out in terms of the original parameters, we have

$$\begin{aligned} \hat{T}_{sys,v} &= T_v - T_Q \sin^2 \Omega + \frac{T_U}{2} \sin 2\Omega + T_{RX,v} + \Delta T_{sys,v}, \\ \hat{T}_{sys,h} &= T_h + T_Q \sin^2 \Omega - \frac{T_U}{2} \sin 2\Omega + T_{RX,h} + \Delta T_{sys,h}, \\ \hat{T}_{sys,U} &= -T_Q \sin 2\Omega + T_U \cos 2\Omega + \Delta T_{sys,U}, \end{aligned} \tag{A.32}$$

and

$$\hat{T}_{sys,I} = T_I + T_{RX,I} + \Delta T_{sys,I}, \tag{A.33}$$

$$\hat{T}_{sys,Q} = T_Q \cos 2\Omega + T_U \sin 2\Omega + T_{RX,Q} + \Delta T_{sys,Q}. \tag{A.34}$$

Now note that $T_{RX,v}$ and $T_{RX,h}$ (and hence also their sum and difference, $T_{RX,I}$ and $T_{RX,Q}$) are operationally estimated and subtracted off as part of the radiometer data calibration. Imperfection in this correction leaves residuals which I call $\Delta T_{RX,v}$ and $\Delta T_{RX,h}$. It is convenient to also define $\Delta T_{RX,I} \equiv \Delta T_{RX,v} + \Delta T_{RX,h}$ and $\Delta T_{RX,Q} \equiv \Delta T_{RX,v} - \Delta T_{RX,h}$. With $T_{RX,v}$, $T_{RX,h}$, $T_{RX,I}$ and $T_{RX,Q}$ subtracted off and leaving only these residuals, we finally have a forward model for the outputs of the rotation, measurement, and calibration processes, which become the inputs to the polarization rotation correction process of [1]. Using a notation similar to [1] for these inputs, where the subscript "a" can be interpreted as referring to temperatures "after" rotation, measurement, and calibration,

$$\hat{T}_{va} = T_v - T_Q \sin^2 \Omega + \frac{T_U}{2} \sin 2\Omega + \Delta T_{RX,v} + \Delta T_{sys,v}, \tag{A.35}$$

$$\hat{T}_{ha} = T_h + T_Q \sin^2 \Omega - \frac{T_U}{2} \sin 2\Omega + \Delta T_{RX,h} + \Delta T_{sys,h}, \tag{A.36}$$

$$\hat{T}'_{Ua} = -T_Q \sin 2\Omega + T_U \cos 2\Omega + \Delta T_{sys,U} \quad . \tag{A.37}$$

As explained in Section 3.2, the measurement and calibration process also add a residual bias, $\Delta T_{RX,U}$, to this last equation, as included in (3.3).

(A.35) and (A.36) are generalizations of equations (12) and (13) in [1]. For convenience, I hereafter use the sum and difference of (A.35) and (A.36), as given in (3.1) and (3.2), respectively.

Appendix B

Derivation of Covariance Matrix, C

Sixteen voltage comprise the operational calibration measurements made by the class of polarimetric radiometer which is considered in Chapter 4. A forward model of these voltages is equation (4.10). Using the information stated in Section 4.4.1 about nine noise variables, this appendix derives the variances and covariances of the sixteen voltages in (4.10). Each column of voltages in (4.10) is independent of the other twelve voltages, so we proceed column by column.

B.1 First Column

Consider the first column of voltages in (4.10). There are just two noise terms, and they are independent of one another since they are from different sources. Therefore $v_{v,C}$ and $v_{h,C}$ are independent. All other relationships have nonzero correlation:

$$Cov(v_{v,C}, v_{p,C}) = Cov(G_{vv}n_1, G_{pv}n_1) = E(G_{vv}G_{pv}n_1^2) \tag{B.1}$$

$$= G_{vv}G_{pv}\frac{(T_C + T_1)^2}{B\tau_c}. \tag{B.2}$$

Similarly,

$$Cov(v_{h,C}, v_{p,C}) = G_{hh}G_{ph}\frac{(T_C + T_2)^2}{B\tau_c}, \tag{B.3}$$

$$Cov(v_{v,C}, v_{m,C}) = G_{vv}G_{mv}\frac{(T_C + T_1)^2}{B\tau_c}, \tag{B.4}$$

$$Cov(v_{h,C}, v_{m,C}) = G_{hh}G_{mh}\frac{(T_C + T_2)^2}{B\tau_c}. \tag{B.5}$$

For the final covariance for this column,

$$\begin{aligned} Cov(v_{p,C}, v_{m,C}) &= E\left((G_{pv}n_1 + G_{ph}n_2)(G_{mv}n_1 + G_{mh}n_2)\right) &\text{(B.6)} \\ &= E\left(G_{pv}G_{mv}n_1^2 + G_{ph}G_{mh}n_2^2\right) &\text{(B.7)} \\ &= G_{pv}G_{mv}\frac{(T_C + T_1)^2}{B\tau_c} + G_{ph}G_{mh}\frac{(T_C + T_2)^2}{B\tau_c}. &\text{(B.8)} \end{aligned}$$

By inspection, the four variances are $G_{vv}^2\frac{(T_C+T_1)^2}{B\tau_c}$, $G_{hh}^2\frac{(T_C+T_2)^2}{B\tau_c}$, $G_{pv}^2\frac{(T_C+T_1)^2}{B\tau_c} + G_{ph}^2\frac{(T_C+T_2)^2}{B\tau_c}$, and $G_{mv}^2\frac{(T_C+T_1)^2}{B\tau_c} + G_{mh}^2\frac{(T_C+T_2)^2}{B\tau_c}$.

B.2 Second Column

The variances and covariances of the second column of voltages in (4.10) are the same as those of the first column, except replacing $(T_C + T_1)$ and $(T_C + T_2)$ with $(T_H + T_1)$ and $(T_H + T_2)$, respectively.

B.3 Third Column

The variances and covariances of the third column are the same as those of the first except replacing $(T_C + T_2)$ with $(T_H + T_2)$.

B.4 Fourth Column

The variances and covariances of the fourth column are quite different because T_{CN} is a component of all three inputs. We first rewrite the calibration inputs (T_C for both the vertical and horizontal channels, $T_{CN}/2$, T_1, and T_2) in terms of electric fields.

The v-channel cold load emits an electric field which we denote c_1. Its second moment (defined as $< c_1^2 >$ where $< \cdot >$ is ensemble average) is T_C (here and hereafter we ignore a constant that converts the product of two electric fields to a brightness temperature). The v-channel amplifier noise is another source, whose equivalent electric field (referred to the input of the first amplifier so that it is on the same level as c_1) is denoted r_1. Its second moment is T_1. Similarly, the h-channel cold load outputs c_2, with second moment T_C, and the h-channel amplifier noise is r_2, whose second moment is T_2.

The correlated calibration source (depicted in Fig. 1 of [2]) emits an electric field n with second moment T_{CN}. When the energy from this source is split between the vertical and horizontal channels, the electric field in each channel is then $n/\sqrt{2}$, whose second moment is $T_{CN}/2$.

The five electric fields just described (c_1, r_1, c_2, r_2, and n) are independent of one another because of their distinct origins. They are all zero-mean normal random variables.

The voltages in the fourth column on the left side of (4.2) are found by summing these electric fields, squaring, integrating, and multiplying by a channel gain,

$$\begin{aligned}
v_{v,CN} &= \frac{G_{vv}}{\tau_c}\int_0^{\tau_c}(c_1 + \frac{n}{\sqrt{2}} + r_1)^2 dt \equiv G_{vv} I, &\text{(B.9)}\\
v_{h,CN} &= \frac{G_{hh}}{\tau_c}\int_0^{\tau_c}(c_2 + \frac{n}{\sqrt{2}} + r_2)^2 dt \equiv G_{hh} J, &\text{(B.10)}\\
v_{p,CN} &= G_{pv} I + G_{ph} J + \frac{G_{pU}}{\tau_c}\int_0^{\tau_c} n^2 dt \equiv G_{pv} I + G_{ph} J + G_{pU} K, &\text{(B.11)}\\
v_{m,CN} &= G_{mv} I + G_{mh} J + G_{mU} K, &\text{(B.12)}
\end{aligned}$$

where the third term in the last two equations arises from the correlation of the inputs to the hybrid coupler.

Because all these voltages are expressed in terms of the I, J, and K defined by these equations, all the variances and covariances can be expressed in terms of the variances and covariances of I, J, and K, which we proceed to determine below.

As shown in [7], I, J, and K can be rewritten as sums of independent samples,

$$I = \frac{1}{N_c}\sum_{i=1}^{N_c}(c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i})^2, \tag{B.13}$$

where $N_c = 2B\tau_c$, B is the sensor bandwidth, and τ_c is the integration time, and similarly with J and K.

B.4.1 Means and Variances of I, J, and K

First, using the independence of each of the N_c samples from one another and of the five electric fields from one another,

$$\begin{aligned}
\langle I \rangle &= \left\langle (c_1 + n/\sqrt{2} + r_1)^2 \right\rangle = \left\langle c_1^2 + n^2/2 + r_1^2 + 2c_1 n/\sqrt{2} + 2c_1 r_1 + 2nr_1/\sqrt{2} \right\rangle \\
&= T_C + T_{CN}/2 + T_1, \qquad \text{(B.14)} \\
\langle J \rangle &= \left\langle (c_2 + n/\sqrt{2} + r_2)^2 \right\rangle = T_C + T_{CN}/2 + T_2, \qquad \text{(B.15)} \\
\langle K \rangle &= \left\langle n^2 \right\rangle = T_{CN}. \qquad \text{(B.16)}
\end{aligned}$$

These means coincide with the final column of the temperature matrix in (4.2), verifying our formulation of the problem in terms of electric fields.

$$Var(I) = \left\langle \frac{1}{N_c} \sum_{i=1}^{N_c} (c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i})^2 \frac{1}{N_c} \sum_{j=1}^{N_c} (c_{1,j} + \frac{n_j}{\sqrt{2}} + r_{1,j})^2 \right\rangle - \langle I \rangle^2 \qquad \text{(B.17)}$$

$$= \frac{1}{N_c^2} \left\langle \sum_{i=1}^{N_c} \sum_{j=1}^{N_c} (c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i})^2 (c_{1,j} + \frac{n_j}{\sqrt{2}} + r_{1,j})^2 \right\rangle - \langle I \rangle^2 . \qquad \text{(B.18)}$$

Separating the expected value operation into terms for which $i \neq j$ and for which $i = j$:

$$\begin{aligned}
Var(I) = \; & \tfrac{1}{N_c^2} \Big(\textstyle\sum_{i=1}^{N_c} \sum_{j=1(\neq i)}^{N_c} \left\langle (c_{1,i} + \tfrac{n_i}{\sqrt{2}} + r_{1,i})^2 (c_{1,j} + \tfrac{n_j}{\sqrt{2}} + r_{1,j})^2 \right\rangle \\
& + \textstyle\sum_{i=1}^{N_c} \left\langle (c_{1,i} + \tfrac{n_i}{\sqrt{2}} + r_{1,i})^4 \right\rangle \Big) - \langle I \rangle^2 . \qquad \text{(B.19)}
\end{aligned}$$

Using the independence of samples i and j, the independence of c_1, n, and r_1 from everything but themselves, and the known fourth moment of zero-mean normal random variables,

$$\begin{aligned}
Var(I) = \; & \tfrac{1}{N_c^2} \Big(\textstyle\sum_{i=1}^{N_c} \left\langle (c_{1,i} + \tfrac{n_i}{\sqrt{2}} + r_{1,i})^2 \right\rangle \sum_{j=1(\neq i)}^{N_c} \left\langle (c_{1,j} + \tfrac{n_j}{\sqrt{2}} + r_{1,j})^2 \right\rangle \\
& + \textstyle\sum_{i=1}^{N_c} 3 \left\langle (c_{1,i} + \tfrac{n_i}{\sqrt{2}} + r_{1,i})^2 \right\rangle^2 \Big) - \langle I \rangle^2 . \qquad \text{(B.20)}
\end{aligned}$$

Using (B.14),

$$Var(I) = \frac{1}{N_c^2} \left(N_c \langle I \rangle (N_c - 1) \langle I \rangle + 3N_c \langle I \rangle^2 \right) - \langle I \rangle^2 = \frac{2}{N_c} \langle I \rangle^2 \qquad \text{(B.21)}$$

and we finally arrive at the ensemble variance

$$Var(I) = \frac{\langle I \rangle^2}{B\tau_c} \text{ and similarly, } Var(J) = \frac{\langle J \rangle^2}{B\tau_c}. \tag{B.22}$$

Also,

$$\begin{aligned}
Var(K) &= \left\langle \frac{1}{N_c}\sum_{i=1}^{N_c} n_i^2 \frac{1}{N_c}\sum_{j=1}^{N_c} n_j^2 \right\rangle - T_{CN}^2 && \text{(B.23)}\\
&= \frac{1}{N_c^2}\left\langle \sum_{i=1}^{N_c}\sum_{j=1}^{N_c} n_i^2 n_j^2 \right\rangle - T_{CN}^2 && \text{(B.24)}\\
&= \frac{1}{N_c^2}\left(\sum_{i=1}^{N_c}\sum_{j=1(\neq i)}^{N_c} \langle n_i^2 n_j^2 \rangle + \sum_{i=1}^{N_c} \langle n_i^4 \rangle \right) - T_{CN}^2 && \text{(B.25)}\\
&= \frac{1}{N_c^2}\left(\sum_{i=1}^{N_c} \langle n_i^2 \rangle \sum_{j=1(\neq i)}^{N_c} \langle n_j^2 \rangle + \sum_{i=1}^{N_c} 3\langle n_i^2 \rangle^2 \right) - T_{CN}^2 && \text{(B.26)}\\
&= \frac{1}{N_c^2}\left(N_c(N_c-1)T_{CN}^2 + 3N_cT_{CN}^2 \right) - T_{CN}^2 && \text{(B.27)}\\
&= \frac{T_{CN}^2}{B\tau_c}. && \text{(B.28)}
\end{aligned}$$

B.4.2 Covariances of I, J, and K

$$\begin{aligned}
Cov(I,J) &= \left\langle \frac{1}{N_c}\sum_{i=1}^{N_c}(c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i})^2 \frac{1}{N_c}\sum_{j=1}^{N_c}(c_{2,j} + \frac{n_j}{\sqrt{2}} + r_{2,j})^2 \right\rangle - \langle I \rangle \langle J \rangle && \text{(B.29)}\\
&= \frac{1}{N_c^2}\left\langle \sum_{i=1}^{N_c}\sum_{j=1}^{N_c}(c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i})^2(c_{2,j} + \frac{n_j}{\sqrt{2}} + r_{2,j})^2 \right\rangle - \langle I \rangle \langle J \rangle . && \text{(B.30)}
\end{aligned}$$

Focus only on the expected value operation, which we separate into terms for which $i \neq j$ and for which $i = j$:

$$\begin{aligned}
&= \sum_{i=1}^{N_c}\sum_{j=1(\neq i)}^{N_c} \left\langle (c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i})^2(c_{2,j} + \frac{n_j}{\sqrt{2}} + r_{2,j})^2 \right\rangle \\
&\quad + \sum_{i=1}^{N_c} \left\langle (c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i})^2(c_{2,i} + \frac{n_i}{\sqrt{2}} + r_{2,i})^2 \right\rangle . && \text{(B.31)}
\end{aligned}$$

Using the independence of samples i and j, the independence of c_1, c_2, r_1, and r_2 from everything but themselves, and the known fourth moment of zero-mean normal random variables,

$$= N_c \langle I \rangle (N_c - 1) \langle J \rangle + N_c \left[\langle I \rangle \ \langle J \rangle + \frac{T_{CN}^2}{2} \right] = N_c^2 \langle I \rangle \ \langle J \rangle + N_c \frac{T_{CN}^2}{2}. \tag{B.32}$$

Putting this back into (B.30),

$$Cov(I,J) = \frac{T_{CN}^2}{4B\tau_c}. \tag{B.33}$$

Similarly we have

$$\begin{aligned}
Cov(I,K) &= \left\langle \frac{1}{N_c} \sum_{i=1}^{N_c} (c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i})^2 \frac{1}{N_c} \sum_{j=1}^{N_c} n_j^2 \right\rangle - \langle I \rangle \ T_{CN} && \text{(B.34)} \\
&= \frac{1}{N_c^2} \left\langle \sum_{i=1}^{N_c} \sum_{j=1}^{N_c} (c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i})^2 n_j^2 \right\rangle - \langle I \rangle \ T_{CN} && \text{(B.35)} \\
&= \frac{1}{N_c^2} \left(\sum_{i=1}^{N_c} \left\langle (c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i})^2 \right\rangle \sum_{j=1(\neq i)}^{N_c} \langle n_j^2 \rangle \right. && \text{(B.36)} \\
&\quad \left. + \sum_{i=1}^{N_c} \left\langle (c_{1,i} + \frac{n_i}{\sqrt{2}} + r_{1,i})^2 n_i^2 \right\rangle \right) - \langle I \rangle \ T_{CN} && \text{(B.37)} \\
&= \frac{1}{N_c^2} \left(N_c \langle I \rangle (N_c - 1) T_{CN} + N_c [T_C T_{CN} + \frac{\langle n^4 \rangle}{2} + T_1 T_{CN}] \right) - \langle I \rangle \ T_{CN} \\
&= \frac{1}{N_c^2} \left(-N_c \langle I \rangle T_{CN} + N_c T_{CN} [T_C + \frac{3}{2} T_{CN} + T_1] \right) && \text{(B.38)} \\
&= -\frac{\langle I \rangle \ T_{CN}}{N_c} + \frac{(\langle I \rangle + T_{CN}) T_{CN}}{N_c} && \text{(B.39)} \\
&= \frac{T_{CN}^2}{2B\tau_c}, && \text{(B.40)}
\end{aligned}$$

and similarly,

$$Cov(J,K) = \frac{T_{CN}^2}{2B\tau_c}. \tag{B.41}$$

B.4.3 Variances of the Voltages

Now that the means, variances, and covariances of I, J, and K are known, we can find the means, variances, and covariances of the voltages:

$$Var(v_{v,CN}) = G_{vv}^2 Var(I) = G_{vv}^2 \frac{\langle I \rangle^2}{B\tau_c}. \tag{B.42}$$

$$Var(v_{h,CN}) = G_{hh}^2 Var(J) = G_{hh}^2 \frac{\langle J \rangle^2}{B\tau_c}. \tag{B.43}$$

$$\begin{aligned} Var(v_{p,CN}) &= Var(G_{pv}I + G_{ph}J + G_{pU}K) && \text{(B.44)} \\ &= Var(G_{pv}I) + Var(G_{ph}J) + Var(G_{pU}K) \\ &\quad +2[G_{pv}G_{ph}Cov(I,J) + G_{pv}G_{pU}Cov(I,K) + G_{ph}G_{pU}Cov(J,K)] \\ &= G_{pv}^2\frac{\langle I\rangle^2}{B\tau_c} + G_{ph}^2\frac{\langle J\rangle^2}{B\tau_c} + G_{pU}^2\frac{T_{CN}^2}{B\tau_c} \\ &\quad +2G_{pv}G_{ph}\frac{T_{CN}^2}{4B\tau_c} + 2G_{pv}G_{pU}\frac{T_{CN}^2}{2B\tau_c} + 2G_{ph}G_{pU}\frac{T_{CN}^2}{2B\tau_c} && \text{(B.45)} \\ &= \frac{G_{pv}^2\langle I\rangle^2 + G_{ph}^2\langle J\rangle^2 + (G_{pU}^2 + G_{pv}G_{ph}/2 + G_{pv}G_{pU} + G_{ph}G_{pU})T_{CN}^2}{B\tau_c}. && \text{(B.46)} \end{aligned}$$

$v_{p,CN}$ is identical to $v_{m,CN}$ – same realizations of noise – except for multiplication by different G_{xx}. Therefore, variances and covariances of the first will be identical to those of the second if we simply replace G_{px} with G_{mx}:

$$Var(v_{m,CN}) = \frac{G_{mv}^2\langle I\rangle^2 + G_{mh}^2\langle J\rangle^2 + (G_{mU}^2 + G_{mv}G_{mh}/2 + G_{mv}G_{mU} + G_{mh}G_{mU})T_{CN}^2}{B\tau_c}. \quad \text{(B.47)}$$

B.4.4 Covariances of the Voltages

The covariances of $v_{v,CN}$ with the other voltages are found as follows:

$$Cov(v_{v,CN}, v_{h,CN}) = G_{vv}G_{hh}Cov(I,J) = G_{vv}G_{hh}\frac{T_{CN}^2}{4B\tau_c}, \quad \text{(B.48)}$$

$$\begin{aligned} Cov(v_{v,CN}, v_{p,CN}) &= Cov(G_{vv}I, G_{pv}I + G_{ph}J + G_{pU}K) && \text{(B.49)} \\ &= G_{vv}[G_{pv}Var(I) + G_{ph}Cov(I,J) + G_{pU}Cov(I,K)] && \text{(B.50)} \\ &= G_{vv}[G_{pv}\frac{\langle I\rangle^2}{B\tau_c} + G_{ph}\frac{T_{CN}^2}{4B\tau_c} + G_{pU}\frac{T_{CN}^2}{2B\tau_c}] && \text{(B.51)} \\ &= G_{vv}\frac{4G_{pv}\langle I\rangle^2 + G_{ph}T_{CN}^2 + 2G_{pU}T_{CN}^2}{4B\tau_c}, && \text{(B.52)} \end{aligned}$$

and similarly,

$$Cov(v_{v,CN}, v_{m,CN}) = G_{vv}\frac{4G_{mv}\langle I\rangle^2 + G_{mh}T_{CN}^2 + 2G_{mU}T_{CN}^2}{4B\tau_c}. \quad \text{(B.53)}$$

Likewise, the covariances of $v_{h,CN}$ with the $v_{p,CN}$ and $v_{m,CN}$ are

$$\begin{aligned}
Cov(v_{h,CN}, v_{p,CN}) &= Cov(G_{hh}J, G_{pv}I + G_{ph}J + G_{pU}K) && \text{(B.54)}\\
&= G_{hh}[G_{pv}Cov(I,J) + G_{ph}Var(J) + G_{pU}Cov(J,K)] && \text{(B.55)}\\
&= G_{hh}[G_{pv}\frac{T_{CN}^2}{4B\tau_c} + G_{ph}\frac{\langle J\rangle^2}{B\tau_c} + G_{pU}\frac{T_{CN}^2}{2B\tau_c}] && \text{(B.56)}\\
&= G_{hh}\frac{G_{pv}T_{CN}^2 + 4G_{ph}\langle J\rangle^2 + 2G_{pU}T_{CN}^2}{4B\tau_c} && \text{(B.57)}
\end{aligned}$$

and

$$Cov(v_{h,CN}, v_{m,CN}) = G_{hh}\frac{G_{mv}T_{CN}^2 + 4G_{mh}\langle J\rangle^2 + 2G_{mU}T_{CN}^2}{4B\tau_c}. \quad \text{(B.58)}$$

The final covariance is longer:

$$\begin{aligned}
Cov(v_{p,CN}, v_{m,CN}) &= Cov(G_{pv}I + G_{ph}J + G_{pU}K, G_{mv}I + G_{mh}J + G_{mU}K)\\
&= G_{pv}G_{mv}Var(I) + G_{ph}G_{mh}Var(J) + G_{pU}G_{mU}Var(K)\\
&\quad +(G_{pv}G_{mh} + G_{ph}G_{mv})Cov(I,J)\\
&\quad +(G_{pv}G_{mU} + G_{pU}G_{mv})Cov(I,K)\\
&\quad +(G_{ph}G_{mU} + G_{pU}G_{mh})Cov(J,K), && \text{(B.59)}
\end{aligned}$$

which expands and then simplifies to

$$\begin{aligned}
Cov(v_{p,CN}, v_{m,CN}) &= \frac{1}{4B\tau_c}[4G_{pv}G_{mv}\langle I\rangle^2 + 4G_{ph}G_{mh}\langle J\rangle^2\\
&\quad +(4G_{pU}G_{mU} + G_{pv}G_{mh} + G_{ph}G_{mv} +\\
&\quad 2G_{pv}G_{mU} + 2G_{pU}G_{mv} + 2G_{ph}G_{mU} + 2G_{pU}G_{mh})T_{CN}^2]. && \text{(B.60)}
\end{aligned}$$

B.5 Summary: entire C

To recapitulate,

$$\mathbf{C} = \begin{bmatrix} C_C & 0 & 0 & 0\\ 0 & C_H & 0 & 0\\ 0 & 0 & C_{CH} & 0\\ 0 & 0 & 0 & C_{CN} \end{bmatrix}, \quad \text{(B.61)}$$

where (note that the following matrices are all symmetric; to make them fit better, only the upper triangular elements are given)

$$C_C = \tfrac{1}{B\tau_c}\times$$

$$\begin{bmatrix} G_{vv}^2(T_C+T_1)^2 & 0 & G_{vv}G_{pv}(T_C+T_1)^2 & G_{vv}G_{mv}(T_C+T_1)^2 \\ & G_{hh}^2(T_C+T_2)^2 & G_{hh}G_{ph}(T_C+T_2)^2 & G_{hh}G_{mh}(T_C+T_2)^2 \\ & & \begin{array}{c} G_{pv}^2(T_C+T_1)^2 \\ +G_{ph}^2(T_C+T_2)^2 \end{array} & \begin{array}{c} G_{pv}G_{mv}(T_C+T_1)^2 \\ +G_{ph}G_{mh}(T_C+T_2)^2 \end{array} \\ & & & \begin{array}{c} G_{mv}^2(T_C+T_1)^2 \\ +G_{mh}^2(T_C+T_2)^2 \end{array} \end{bmatrix} \tag{B.62}$$

$$C_H = \tfrac{1}{B\tau_c}\times$$

$$\begin{bmatrix} G_{vv}^2(T_H+T_1)^2 & 0 & G_{vv}G_{pv}(T_H+T_1)^2 & G_{vv}G_{mv}(T_H+T_1)^2 \\ & G_{hh}^2(T_H+T_2)^2 & G_{hh}G_{ph}(T_H+T_2)^2 & G_{hh}G_{mh}(T_H+T_2)^2 \\ & & \begin{array}{c} G_{pv}^2(T_H+T_1)^2 \\ +G_{ph}^2(T_H+T_2)^2 \end{array} & \begin{array}{c} G_{pv}G_{mv}(T_H+T_1)^2 \\ +G_{ph}G_{mh}(T_H+T_2)^2 \end{array} \\ & & & \begin{array}{c} G_{mv}^2(T_H+T_1)^2 \\ +G_{mh}^2(T_H+T_2)^2 \end{array} \end{bmatrix} \tag{B.63}$$

$$C_{CH} = \tfrac{1}{B\tau_c}\times$$

$$\begin{bmatrix} G_{vv}^2(T_C+T_1)^2 & 0 & G_{vv}G_{pv}(T_C+T_1)^2 & G_{vv}G_{mv}(T_C+T_1)^2 \\ & G_{hh}^2(T_H+T_2)^2 & G_{hh}G_{ph}(T_H+T_2)^2 & G_{hh}G_{mh}(T_H+T_2)^2 \\ & & \begin{array}{c} G_{pv}^2(T_C+T_1)^2 \\ +G_{ph}^2(T_H+T_2)^2 \end{array} & \begin{array}{c} G_{pv}G_{mv}(T_C+T_1)^2 \\ +G_{ph}G_{mh}(T_H+T_2)^2 \end{array} \\ & & & \begin{array}{c} G_{mv}^2(T_C+T_1)^2 \\ +G_{mh}^2(T_H+T_2)^2 \end{array} \end{bmatrix} \tag{B.64}$$

and

$$C_{CN} = \tfrac{1}{B\tau_c}\times$$

$$\begin{bmatrix} G_{vv}^2\left\langle I\right\rangle^2 & G_{vv}G_{hh}\tfrac{T_{CN}^2}{4} & G_{vv}\tfrac{4G_{pv}\left\langle I\right\rangle^2+G_{ph}T_{CN}^2+2G_{pU}T_{CN}^2}{4} & G_{vv}\tfrac{4G_{mv}\left\langle I\right\rangle^2+G_{mh}T_{CN}^2+2G_{mU}T_{CN}^2}{4} \\ & G_{hh}^2\left\langle J\right\rangle^2 & G_{hh}\tfrac{G_{pv}T_{CN}^2+4G_{ph}\left\langle J\right\rangle^2+2G_{pU}T_{CN}^2}{4} & G_{hh}\tfrac{G_{mv}T_{CN}^2+4G_{mh}\left\langle J\right\rangle^2+2G_{mU}T_{CN}^2}{4} \\ & & P & Y \\ & & & M \end{bmatrix}, \tag{B.65}$$

where

$$\begin{aligned} P &\equiv G_{pv}^2\left\langle I\right\rangle^2 + G_{ph}^2\left\langle J\right\rangle^2 + \left(G_{pU}^2 + G_{pv}G_{ph}/2 + G_{pv}G_{pU} + G_{ph}G_{pU}\right)T_{CN}^2, \\ M &\equiv G_{mv}^2\left\langle I\right\rangle^2 + G_{mh}^2\left\langle J\right\rangle^2 + \left(G_{mU}^2 + G_{mv}G_{mh}/2 + G_{mv}G_{mU} + G_{mh}G_{mU}\right)T_{CN}^2, \\ Y &\equiv G_{pv}G_{mv}\left\langle I\right\rangle^2 + G_{ph}G_{mh}\left\langle J\right\rangle^2 + \Bigg(G_{pU}G_{mU} + \frac{G_{pv}G_{mh}}{4} + \frac{G_{ph}G_{mv}}{4} \\ &\quad + \frac{G_{pv}G_{mU}}{2} + \frac{G_{pU}G_{mv}}{2} + \frac{G_{ph}G_{mU}}{2} + \frac{G_{pU}G_{mh}}{2}\Bigg)T_{CN}^2, \end{aligned}$$

and where $\langle I\rangle$ and $\langle J\rangle$ are given in (B.14) and (B.15), respectively.

Appendix C

Derivation of Constraint Equations

Chapter 4 is concerned with the estimation of eight gains (among other parameters) which characterize a certain class of polarimetric radiometer. This appendix contains the derivation of equations constraining those eight gains.

C.1 Finding V_2 (eigenvectors of C for which $\lambda = 0$)

Due to the block diagonal structure of the covariance matrix C of calibration measurements (see Appendix B), C has four eigenvectors of the form $[abcd000000000000]^T$ where $[abcd]^T$ is an eigenvector of C_C, four of the form $[0000abcd00000000]^T$ where $[abcd]^T$ is an eigenvector of C_H, four of the form $[00000000abcd0000]^T$ where $[abcd]^T$ is an eigenvector of C_{CH}, and four of the form $[000000000000efgh]^T$ where $[efgh]^T$ is an eigenvector of C_{CN}. C_C, C_H, and C_{CH} each have two eigenvalues (λ) equal to zero while C_{CN} has only one. This is easily confirmed numerically; theoretically, it is because the first three columns of (4.10) have two noise sources each while the last has three.

C.1.1 Eigenvectors of C_C, C_H, and C_{CH} for which $\lambda = 0$

C_C, C_H, and C_{CH} can all be written in the abbreviated forms

$$\frac{1}{B\tau_c}\begin{bmatrix} v^2T & 0 & vpT & vmT \\ 0 & h^2U & hqU & hnU \\ vpT & hqU & p^2T + q^2U & pmT + qnU \\ vmT & hnU & pmT + qnU & m^2T + n^2U \end{bmatrix}, \tag{C.1}$$

where the only difference between C_C, C_H, and C_{CH} is whether T and U are defined using T_C or T_H. The eigenvectors of this matrix are found using the defining equation of an eigenvector with $\lambda = 0$, explicitly

$$\begin{bmatrix} v^2T & 0 & vpT & vmT \\ 0 & h^2U & hqU & hnU \\ vpT & hqU & p^2T + q^2U & pmT + qnU \\ vmT & hnU & pmT + qnU & m^2T + n^2U \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{C.2}$$

From the first two of the four equations in (C.2),

$$a = -\frac{cp + dm}{v} \text{ and } b = -\frac{cq + dn}{h}. \tag{C.3}$$

Using these to substitute for a and b in the third and fourth equations in (C.2), these equations reduce to

$$0c + 0d = 0 \text{ and } 0c + 0d = 0. \tag{C.4}$$

Any c and d will satisfy these equations. We choose simple but nontrivial values: $c = 1, d = 0$ and $c = 0, d = 1$. Then, using (C.3), the eigenvectors are

$$\begin{bmatrix} -p/v \\ -q/h \\ 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} -m/v \\ -n/h \\ 0 \\ 1 \end{bmatrix}. \tag{C.5}$$

Eigenvectors can be scaled arbitrarily. Scaling by $-vh$ and undoing the abbreviations, the eigenvectors with $\lambda = 0$ are[1]:

$$\begin{bmatrix} G_{hh}G_{pv} \\ G_{vv}G_{ph} \\ -G_{vv}G_{hh} \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} G_{hh}G_{mv} \\ G_{vv}G_{mh} \\ 0 \\ -G_{vv}G_{hh} \end{bmatrix}. \tag{C.6}$$

---

[1] To within about 10%, the eigenvectors are $\approx G_{vv}G_{hh}[\frac{1}{2}\ \frac{1}{2}\ -1\ 0]^T$ and $G_{vv}G_{hh}[\frac{1}{2}\ \frac{1}{2}\ 0\ -1]$.

Note that these are *exactly the same* (though they can be scaled arbitrarily and independently) for C_C as for C_H and for C_{CH}, since they do not depend on T and U[2][3].

C.1.2 Eigenvector of C_{CN} for which $\lambda = 0$

Abbreviate the eigenvector problem for C_{CN} as

$$\begin{bmatrix} V & X & A & B \\ X & H & C & D \\ A & C & P & Y \\ B & D & Y & M \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{C.7}$$

The procedure is similar to that used above but now with an X instead of zeros. We solve the first equation for a, giving $a = \frac{-bX-cA-dB}{V}$. Using this to substitute for a in the second equation, then solving that equation for b, gives

$$\begin{aligned} b &= \frac{-aX - cC - dD}{H} && \text{(C.8)} \\ &= \frac{\frac{bX+cA+dB}{V}X - cC - dD}{H} && \text{(C.9)} \\ &= \frac{(bX + cA + dB)X - cCV - dDV}{VH}, && \text{(C.10)} \end{aligned}$$

$$\begin{aligned} b(1 - X^2/VH) &= \frac{cAX + dBX - cCV - dDV}{VH}, && \text{(C.11)} \\ b(VH - X^2) &= cAX + dBX - cCV - dDV, && \text{(C.12)} \\ b &= \frac{c(AX - CV) + d(BX - DV)}{VH - X^2}. && \text{(C.13)} \end{aligned}$$

---

[2] When V_2 is found numerically, there are usually four nonzero entries rather than the three predicted by the above derivation. Both answers are correct, as can be verified by showing that in either case, $C_C V_2$ equals a zero 16x7 matrix . Why is this so? Since there are two variables that can be chosen arbitrarily, their span is a 2-D hyperplane in 4-D space. Any two vectors whose span is that same plane work, and that is why both answers are correct. The analytical answer just happens to be formed more simply, with a couple of zero elements. In (C.4), any c and d could work. In fact if we put the c and d that a numerical calculation returns into the formulas above for a and b, they match the a and b that the numerical calculation returns to at least 14 significant digits.

[3] In summary, the eigenvectors of C_C, C_H, and C_{CH} whose $\lambda = 0$ are found as follows. First choose c and d arbitrarily, as long as they are not both zero. If $c = vh$ and $d = 0$ are chosen, then one eigenvector each of C_C, C_H, and C_{CH} is given (in terms of the original parameters) by the first vector in (C.6)–otherwise use the more general formulas above to find a and b. Next, choose another c and d arbitrarily, as long as they are not both zero and are not simply a scaled version of the first choice. If $c = 0$ and $d = vh$ are chosen, then the second eigenvector each of C_C, C_H, and C_{CH} is given (in terms of the original parameters) by the second vector in (C.6) – otherwise use more general formulas above to find a and b.

Using this, we can express a in terms of c and d as:

$$\begin{aligned}
a &= \frac{-\frac{c(AX-CV)+d(BX-DV)}{VH-X^2}X - cA - dB}{V} && \text{(C.14)}\\
&= \frac{-c(AX-CV)X - d(BX-DV)X}{V(VH-X^2)} - c\frac{A}{V} - d\frac{B}{V} && \text{(C.15)}\\
&= \frac{c[(AX-CV)X + A(VH-X^2)] + d[(BX-DV)X + B(VH-X^2)]}{-V(VH-X^2)}. && \text{(C.16)}
\end{aligned}$$

Using both the above results to substitute for a and b in the third equation, then solving that equation for c, gives

$$\begin{aligned}
c &= \frac{-aA - bC - dY}{P} && \text{(C.17)}\\
&= \frac{cA[(AX-CV)X + A(VH-X^2)] + dA[(BX-DV)X + B(VH-X^2)]}{PV(VH-X^2)}\\
&\quad + \frac{-cC(AX-CV) - dC(BX-DV)}{P(VH-X^2)} + \frac{-dY}{P} && \text{(C.18)}\\
&= \frac{c[(AX-CV)^2 + A^2(VH-X^2)]}{PV(VH-X^2)}\\
&\quad + \frac{d[(BX-DV)(AX-CV) + AB(VH-X^2)]}{PV(VH-X^2)} + \frac{-dY}{P}, && \text{(C.19)}
\end{aligned}$$

$$\begin{aligned}
c &\left[1 - \frac{(AX-CV)^2 + A^2(VH-X^2)}{PV(VH-X^2)}\right]\\
&= d\frac{(BX-DV)(AX-CV) + AB(VH-X^2)}{PV(VH-X^2)} + \frac{-dY}{P}, && \text{(C.20)}\\
c &\left[PV(VH-X^2) - (AX-CV)^2 - A^2(VH-X^2)\right]\\
&= d\left[(BX-DV)(AX-CV) + AB(VH-X^2) - YV(VH-X^2)\right], && \text{(C.21)}
\end{aligned}$$

$$\begin{aligned}
c &= d\frac{(AB-YV)(VH-X^2) + (BX-DV)(AX-CV)}{(PV-A^2)(VH-X^2) - (AX-CV)^2} && \text{(C.22)}\\
c &= -d\frac{ABH - YVH + YX^2 - BCX - ADX + CDV}{A^2H - PVH + PX^2 - 2ACX + C^2V} && \text{(C.23)}\\
c &= -d\left(\frac{Y(VH-X^2) - ABH + (AD+BC)X - CDV}{P(VH-X^2) - A^2H + 2ACX - C^2V}\right) && \text{(C.24)}\\
c &= -d\left(\frac{Y(VH-X^2) + D(AX-CV) + B(CX-AH)}{P(VH-X^2) + C(AX-CV) + A(CX-AH)}\right)\\
&\equiv -dk_3. && \text{(C.25)}
\end{aligned}$$

If we put these expressions for a, b, and c into the fourth equation, it gives $0d = 0$ which is satisfied by any d. Rewriting a and b in terms of d only, not c and d,

$$\begin{aligned} b &= -d\frac{k_3(AX - CV) - (BX - DV)}{VH - X^2} \\ &\equiv -dk_2, \qquad \text{(C.26)} \\ a &= d\left(\frac{k_2X + k_3A - B}{V}\right) \\ &\equiv dk_1, \qquad \text{(C.27)} \end{aligned}$$

and the eigenvector is

$$-d\begin{bmatrix} -k_1 \\ k_2 \\ k_3 \\ -1 \end{bmatrix} = -d\begin{bmatrix} \frac{-k_2X - k_3A + B}{V} \\ \frac{k_3(AX-CV)-(BX-DV)}{VH-X^2} \\ k_3 \\ -1 \end{bmatrix} \tag{C.28}$$

$$= -d\begin{bmatrix} -\left(\frac{Y(VH-X^2)+D(AX-CV)+B(CX-AH)}{P(VH-X^2)+C(AX-CV)+A(CX-AH)} \cdot \frac{AX-CV}{VH-X^2} - \frac{BX-DV}{VH-X^2}\right)\frac{X}{V} \\ -\frac{Y(VH-X^2)+D(AX-CV)+B(CX-AH)}{P(VH-X^2)+C(AX-CV)+A(CX-AH)} \cdot \frac{A}{V} + \frac{B}{V} \\ \\ \frac{Y(VH-X^2)+D(AX-CV)+B(CX-AH)}{P(VH-X^2)+C(AX-CV)+A(CX-AH)} \cdot \frac{AX-CV}{VH-X^2} - \frac{BX-DV}{VH-X^2} \\ \\ \frac{Y(VH-X^2)+D(AX-CV)+B(CX-AH)}{P(VH-X^2)+C(AX-CV)+A(CX-AH)} \\ \\ -1 \end{bmatrix}. \tag{C.29}$$

Choosing d to be the denominator of k_3 and scaling the eigenvector by -1 simplifies the eigenvector to

$$\begin{bmatrix} -\left(\frac{Y(VH-X^2)+D(AX-CV)+B(CX-AH)}{1} \cdot \frac{AX-CV}{VH-X^2} - \frac{d(BX-DV)}{VH-X^2}\right)\frac{X}{V} \\ -\frac{Y(VH-X^2)+D(AX-CV)+B(CX-AH)}{1} \cdot \frac{A}{V} + \frac{dB}{V} \\ \\ \frac{Y(VH-X^2)+D(AX-CV)+B(CX-AH)}{1} \cdot \frac{AX-CV}{VH-X^2} - \frac{d(BX-DV)}{VH-X^2} \\ \\ Y(VH - X^2) + D(AX - CV) + B(CX - AH) \\ \\ -[P(VH - X^2) + C(AX - CV) + A(CX - AH)] \end{bmatrix}. \tag{C.30}$$

An expression in the first element of this eigenvector, namely

$$\left[\frac{\left[Y(VH-X^2)+D(AX-CV)+B(CX-AH)\right](AX-CV)-d(BX-DV)}{VH-X^2}\right], \tag{C.31}$$

reduces to $Y(AX-CV)-P(BX-DV)+A(BC-AD)$. Other terms also simplify, leaving the eigenvector

$$\begin{bmatrix} -\left(Y(AX-CV)-P(BX-DV)+A(BC-AD)\right)\frac{X}{V} \\ +\frac{(BP-AY)(VH-X^2)+(BC-AD)(AX-CV)}{V} \\ \\ Y(AX-CV)-P(BX-DV)+A(BC-AD) \\ \\ Y(VH-X^2)+D(AX-CV)+B(CX-AH) \\ \\ -[P(VH-X^2)+C(AX-CV)+A(CX-AH)] \end{bmatrix}, \tag{C.32}$$

which then simplifies to

$$\begin{bmatrix} P(BH-DX)+Y(CX-AH)-C(BC-AD) \\ \\ Y(AX-CV)-P(BX-DV)+A(BC-AD) \\ \\ D(AX-CV)+B(CX-AH)+Y(VH-X^2) \\ \\ -C(AX-CV)-A(CX-AH)-P(VH-X^2) \end{bmatrix}. \tag{C.33}$$

This was verified numerically. This simplifies when put in terms of the original parameters. A $G_{vv}G_{hh}$ can be factored out of the entire eigenvector; the remaining G_{vv} and/or G_{hh} can be factored out of each element, leaving

$$G_{vv}G_{hh}\begin{bmatrix} G_{hh}(...) \\ G_{vv}(...) \\ G_{vv}G_{hh}(...) \\ G_{vv}G_{hh}(...) \end{bmatrix}, \tag{C.34}$$

where the factors in parentheses have no G_{vv} or G_{hh}. Computer algebra software was then used to simplify this vector further, resulting in Eigenvector of C_{CN} with $\lambda = 0$ is[4]

$$\begin{bmatrix} \frac{G_{pv}G_{mU}-G_{pU}G_{mv}}{G_{vv}} \\ \frac{G_{ph}G_{mU}-G_{pU}G_{mh}}{G_{hh}} \\ -G_{mU} \\ G_{pU} \end{bmatrix}. \tag{C.36}$$

C.2 Forming Constraint Equations from V_2

For fixed $\mathbf{v}$ and unknown $\mathbf{m}$, the constraint is

$$V_2^T(\mathbf{m})\mathbf{v} = V_2^T(\mathbf{m})\mathbf{g}(\mathbf{m}), \tag{C.37}$$

where we explicitly show that V_2 is formed from the unknown $\mathbf{m}$ (through C).

C.2.1 First Six Constraint Equations

Having found the eigenvectors that form V_2 in section C.1, we can now write the first two equations of (C.37) as

$$\begin{aligned} G_{hh}G_{pv}v_{v,C} + G_{vv}G_{ph}v_{h,C} - G_{vv}G_{hh}v_{p,C} = \\ G_{hh}G_{pv}G_{vv}(T_C+T_1) + G_{vv}G_{ph}G_{hh}(T_C+T_2) \\ -G_{vv}G_{hh}[G_{pv}(T_C+T_1) + G_{ph}(T_C+T_2)], \quad \text{(C.38)} \\ G_{hh}G_{mv}v_{v,C} + G_{vv}G_{mh}v_{h,C} - G_{vv}G_{hh}v_{m,C} = \\ G_{hh}G_{mv}G_{vv}(T_C+T_1) + G_{vv}G_{mh}G_{hh}(T_C+T_2) \\ -G_{vv}G_{hh}[G_{mv}(T_C+T_1) + G_{mh}(T_C+T_2)]. \quad \text{(C.39)} \end{aligned}$$

---

[4] Note that when appropriately scaled, this is approximately (i.e., within about 10% of being)

$$G_{pU}\begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}. \tag{C.35}$$

The right sides cancel themselves out, leaving

$$G_{hh}G_{pv}v_{v,C} + G_{vv}G_{ph}v_{h,C} - G_{vv}G_{hh}v_{p,C} = 0, \tag{C.40}$$

$$G_{hh}G_{mv}v_{v,C} + G_{vv}G_{mh}v_{h,C} - G_{vv}G_{hh}v_{m,C} = 0, \tag{C.41}$$

or simply

$$G_{hh}G_{pv}v_{v,C} + G_{vv}G_{ph}v_{h,C} = G_{vv}G_{hh}v_{p,C}, \tag{C.42}$$

$$G_{hh}G_{mv}v_{v,C} + G_{vv}G_{mh}v_{h,C} = G_{vv}G_{hh}v_{m,C}. \tag{C.43}$$

For the next two constraint equations, corresponding to eigenvectors of C_H, the procedure is very similar, leading to the constraint equations

$$G_{hh}G_{pv}v_{v,H} + G_{vv}G_{ph}v_{h,H} = G_{vv}G_{hh}v_{p,H}, \tag{C.44}$$

$$G_{hh}G_{mv}v_{v,H} + G_{vv}G_{mh}v_{h,H} = G_{vv}G_{hh}v_{m,H}. \tag{C.45}$$

Using the eigenvectors of C_{CH}, the fifth and sixth equations of (C.37) become

$$\begin{aligned} G_{hh}G_{pv}v_{v,CH} + G_{vv}G_{ph}v_{h,CH} - G_{vv}G_{hh}v_{p,CH} = \\ G_{hh}G_{pv}G_{vv}(T_C + T_1) + G_{vv}G_{ph}G_{hh}(T_H + T_2) \\ -G_{vv}G_{hh}[G_{pv}(T_C + T_1) + G_{ph}(T_H + T_2)], \end{aligned} \tag{C.46}$$

$$\begin{aligned} G_{hh}G_{mv}v_{v,CH} + G_{vv}G_{mh}v_{h,CH} - G_{vv}G_{hh}v_{m,CH} = \\ G_{hh}G_{mv}G_{vv}(T_C + T_1) + G_{vv}G_{mh}G_{hh}(T_H + T_2) \\ -G_{vv}G_{hh}[G_{mv}(T_C + T_1) + G_{mh}(T_H + T_2)], \end{aligned} \tag{C.47}$$

whose right sides still cancel themselves out. Therefore our first six constraint equations, in six unknowns, are (ordered differently for future convenience)

$$G_{hh}G_{pv}v_{v,C} + G_{vv}G_{ph}v_{h,C} = G_{vv}G_{hh}v_{p,C}, \tag{C.48}$$

$$G_{hh}G_{pv}v_{v,H} + G_{vv}G_{ph}v_{h,H} = G_{vv}G_{hh}v_{p,H}, \tag{C.49}$$

$$G_{hh}G_{pv}v_{v,CH} + G_{vv}G_{ph}v_{h,CH} = G_{vv}G_{hh}v_{p,CH}, \tag{C.50}$$

$$G_{hh}G_{mv}v_{v,C} + G_{vv}G_{mh}v_{h,C} = G_{vv}G_{hh}v_{m,C}, \tag{C.51}$$

$$G_{hh}G_{mv}v_{v,H} + G_{vv}G_{mh}v_{h,H} = G_{vv}G_{hh}v_{m,H}, \tag{C.52}$$

$$G_{hh}G_{mv}v_{v,CH} + G_{vv}G_{mh}v_{h,CH} = G_{vv}G_{hh}v_{m,CH}. \tag{C.53}$$

Multiplying (C.48) by $v_{v,H}/v_{v,C}$ and subtracting (C.49) leaves

$$G_{ph}(v_{h,C}v_{v,H}/v_{v,C} - v_{h,H}) \quad = \quad G_{hh}(v_{p,C}v_{v,H}/v_{v,C} - v_{p,H}). \tag{C.54}$$

Similarly, from (C.48) and (C.50) we obtain

$$G_{ph}(v_{h,C}v_{v,CH}/v_{v,C} - v_{h,CH}) \quad = \quad G_{hh}(v_{p,C}v_{v,CH}/v_{v,C} - v_{p,CH}). \tag{C.55}$$

Solve (C.55) for G_{hh}

$$G_{hh} = G_{ph}\frac{(v_{h,C}v_{v,CH}/v_{v,C} - v_{h,CH})}{(v_{p,C}v_{v,CH}/v_{v,C} - v_{p,CH})}, \tag{C.56}$$

then substitute it into (C.54) to obtain

$$G_{ph}(v_{h,C}v_{v,H}/v_{v,C} - v_{h,H}) = G_{ph}\frac{(v_{h,C}v_{v,CH}/v_{v,C} - v_{h,CH})}{(v_{p,C}v_{v,CH}/v_{v,C} - v_{p,CH})}(v_{p,C}v_{v,H}/v_{v,C} - v_{p,H}). \tag{C.57}$$

This last equation cannot be solved for G_{ph} – it merely reveals redundancy in the data and that (C.54) and (C.55) are redundant constraint equations.

Similarly, from (C.51) and (C.52) we obtain

$$G_{mh}(v_{h,C}v_{v,H}/v_{v,C} - v_{h,H}) \quad = \quad G_{hh}(v_{m,C}v_{v,H}/v_{v,C} - v_{m,H}) \tag{C.58}$$

or an expression that is numerically equivalent from (C.51) and (C.53).

A similar procedure produces constraint equations for G_{pv} and G_{mv} in terms of G_{vv}. In summary we end up with the four constraint equations

$$G_{pv} = G_{vv}\frac{(v_{p,C}v_{h,H} - v_{h,C}v_{p,H})}{(v_{v,C}v_{h,H} - v_{h,C}v_{v,H})},\ G_{ph} = G_{hh}\frac{(v_{p,C}v_{v,H} - v_{v,C}v_{p,H})}{(v_{h,C}v_{v,H} - v_{v,C}v_{h,H})}, \tag{C.59}$$

$$G_{mv} = G_{vv}\frac{(v_{m,C}v_{h,H} - v_{h,C}v_{m,H})}{(v_{v,C}v_{h,H} - v_{h,C}v_{v,H})},\ G_{mh} = G_{hh}\frac{(v_{m,C}v_{v,H} - v_{v,C}v_{m,H})}{(v_{h,C}v_{v,H} - v_{v,C}v_{h,H})}, \tag{C.60}$$

or two alternative (but numerically equivalent) sets obtained by finding, for example, the equation relating G_{ph} and G_{hh} from (C.48) and (C.50) or from (C.49) and (C.50). So of (C.48) through (C.50), any one is completely redundant and the same for (C.51) through (C.53). Hence, rather than six equations in six unknowns we only have four equations in six unknowns.

C.2.2 Last Constraint Equation

Using the last eigenvector with $\lambda = 0$, as derived in section C.1, the final equation from (C.37) becomes

$$\begin{aligned}
&\frac{G_{pv}G_{mU} - G_{pU}G_{mv}}{G_{vv}} v_{v,CN} + \frac{G_{ph}G_{mU} - G_{pU}G_{mh}}{G_{hh}} v_{h,CN} - G_{mU}v_{p,CN} + G_{pU}v_{m,CN} \\
&= \frac{G_{pv}G_{mU} - G_{pU}G_{mv}}{G_{vv}} G_{vv} \langle I \rangle + \frac{G_{ph}G_{mU} - G_{pU}G_{mh}}{G_{hh}} G_{hh} \langle J \rangle \\
&\quad - G_{mU}(G_{pv} \langle I \rangle + G_{ph} \langle J \rangle + G_{pU}T_{CN}) + G_{pU}(G_{mv} \langle I \rangle + G_{mh} \langle J \rangle + G_{mU}T_{CN}).
\end{aligned} \tag{C.61}$$

The entire right side cancels itself out. Then multiplying both sides by $G_{vv}G_{hh}$,

$$\begin{aligned}
&(G_{pv}G_{mU} - G_{pU}G_{mv})G_{hh}v_{v,CN} + (G_{ph}G_{mU} - G_{pU}G_{mh})G_{vv}v_{h,CN} \\
&\qquad - G_{mU}G_{vv}G_{hh}v_{p,CN} + G_{pU}G_{vv}G_{hh}v_{m,CN} = 0.
\end{aligned} \tag{C.62}$$

Solving this for G_{mU},

$$G_{mU} = G_{pU} \cdot \frac{G_{mv}G_{hh}v_{v,CN} + G_{mh}G_{vv}v_{h,CN} - G_{vv}G_{hh}v_{m,CN}}{G_{pv}G_{hh}v_{v,CN} + G_{ph}G_{vv}v_{h,CN} - G_{vv}G_{hh}v_{p,CN}}. \tag{C.63}$$

Appendix D

Constraints on Hardware Parameters

This appendix contains derivations of equations constraining the hardware parameters defined and used in Chapter 4. The first section is concerned with a hybrid coupler scattering parameter, s. The second section is concerned with detector sensitivities.

D.1 Derivation of Equation for s, (4.34)

By replacing the G_{xx} in (4.21)–(4.23) with their definitions in terms of radiometer hardware parameters, the constraint equations become (after canceling out common factors on both sides and simplifying the last equation)

$$c_p s^2 = c_v \frac{(v_{p,C}v_{h,H} - v_{h,C}v_{p,H})}{(v_{v,C}v_{h,H} - v_{h,C}v_{v,H})}, \tag{D.1}$$

$$c_p(1-s^2) = c_h \frac{(v_{p,C}v_{v,H} - v_{v,C}v_{p,H})}{(v_{h,C}v_{v,H} - v_{v,C}v_{h,H})}, \tag{D.2}$$

$$c_m(1-s^2) = c_v \frac{(v_{m,C}v_{h,H} - v_{h,C}v_{m,H})}{(v_{v,C}v_{h,H} - v_{h,C}v_{v,H})}, \tag{D.3}$$

$$c_m s^2 = c_h \frac{(v_{m,C}v_{v,H} - v_{v,C}v_{m,H})}{(v_{h,C}v_{v,H} - v_{v,C}v_{h,H})}, \tag{D.4}$$

$$c_v c_h (c_p v_{m,CN} + c_m v_{p,CN}) = c_p c_m (c_v v_{h,CN} + c_h v_{v,CN}), \tag{D.5}$$

which are five equations in five unknowns. The hardware parameters that cancel out, G_1, G_2, and α_e, are unconstrained, as are T_1 and T_2. It is convenient to define abbreviations for various combinations of voltages:

$$
\begin{aligned}
A &\equiv v_{p,C}v_{h,H} - v_{h,C}v_{p,H}, && \text{(D.6)}\\
B &\equiv v_{p,C}v_{v,H} - v_{v,C}v_{p,H}, && \text{(D.7)}\\
C &\equiv v_{m,C}v_{h,H} - v_{h,C}v_{m,H}, && \text{(D.8)}\\
D &\equiv v_{m,C}v_{v,H} - v_{v,C}v_{m,H}, && \text{(D.9)}\\
E &\equiv v_{v,C}v_{h,H} - v_{h,C}v_{v,H}. && \text{(D.10)}
\end{aligned}
$$

Using these abbreviations, (D.1) through (D.5) can be rewritten as

$$
\begin{aligned}
c_p &= c_v\frac{A}{Es^2}, && \text{(D.11)}\\
c_p &= c_h\frac{B}{-E(1-s^2)}, && \text{(D.12)}\\
c_m &= c_v\frac{C}{E(1-s^2)}, && \text{(D.13)}\\
c_m &= c_h\frac{D}{-Es^2}, && \text{(D.14)}\\
c_v c_h(c_p v_{m,CN} + c_m v_{p,CN}) &= c_p c_m(c_v v_{h,CN} + c_h v_{v,CN}). && \text{(D.15)}
\end{aligned}
$$

Equations (D.11) and (D.14) are already solved for c_p and c_m. Using these to substitute for c_p and c_m in the remaining three constraints yields

$$
\begin{aligned}
c_v\frac{A}{s^2} &= c_h\frac{B}{-(1-s^2)}, && \text{(D.16)}\\
c_h\frac{D}{-s^2} &= c_v\frac{C}{(1-s^2)}, && \text{(D.17)}\\
c_v A v_{m,CN} - c_h D v_{p,CN} &= -\frac{AD}{Es^2}(c_v v_{h,CN} + c_h v_{v,CN}). && \text{(D.18)}
\end{aligned}
$$

Solving the second of these for c_h and substituting this into the remaining two constraints yields

$$
\begin{aligned}
AD(1-s^2)^2 &= BCs^4, && \text{(D.19)}\\
Av_{m,CN} + \frac{Cs^2}{(1-s^2)}v_{p,CN} &= -\frac{AD}{Es^2}(v_{h,CN} - \frac{Cs^2}{D(1-s^2)}v_{v,CN}). && \text{(D.20)}
\end{aligned}
$$

The fact that c_v cancels out of both of these indicates that c_v is unconstrained, like G_1, G_2, α_e, T_1, and T_2. (If we had solved the equations differently, any one of the c_x

could be the unconstrained one.) Both of the above equations can be rearranged as quadratic equations in s^2, namely

$$(AD - BC)s^4 - 2ADs^2 + AD = 0, \tag{D.21}$$

$$(AEv_{m,CN} - CEv_{p,CN})s^4 + (-AEv_{m,CN} + ADv_{h,CN} + ACv_{v,CN})s^2 - ADv_{h,CN} = 0. \tag{D.22}$$

The first is solved using the quadratic formula; its positive solution is

$$s^2 = \frac{AD + \sqrt{ADBC}}{AD - BC}. \tag{D.23}$$

The second is also solved using the quadratic formula; its positive solution turns out to be the same[1]. Therefore the last two constraint equations are redundant of one another. From (D.23) we then obtain (4.34).

D.2 Derivation of Equations Relating c_h, c_p, and c_m to c_v, (4.35) through (4.37)

Combining (D.13) and (D.14) to solve for c_h in terms of c_v yields

$$c_h = -c_v \frac{Cs^2}{D(1 - s^2)}. \tag{D.25}$$

Joining this with (D.11) and (D.13) gives c_h, c_p, and c_m in terms of c_v:

$$\begin{aligned} c_h &= -c_v \frac{Cs^2}{D(1 - s^2)}, && \text{(D.26)} \\ c_p &= c_v \frac{A}{Es^2}, && \text{(D.27)} \\ c_m &= c_v \frac{C}{E(1 - s^2)}. && \text{(D.28)} \end{aligned}$$

Substituting for s^2 using (D.23) yields (4.35) through (4.37).

---

[1] We have shown this numerically. These solutions are also numerically the same as the solution to an equation derived by another route, namely

$$\begin{aligned} &(AEv_{m,CN} - CEv_{p,CN})s^4 + (-2AEv_{m,CN} + CEv_{p,CN} + BCv_{h,CN} + ACv_{v,CN})s^2 \\ &\qquad + AEv_{m,CN} - ACv_{v,CN} = 0. \end{aligned} \tag{D.24}$$

Appendix E

Description of Chapter 5 Datasets

This appendix describes datasets used within Chapter 5. These are air temperature datasets, a dataset of satellite microwave brightness temperatures, and a combined dataset.

E.1 Air Temperature Data and Station Geography

The air temperature data T used in this work are available at dss.ucar.edu, within Dataset 464.0. Our 2007-2008 data is from a more frequently updated site, http://www.antarctica.ac.uk/met/metlog/cui.html.

The first three locations used in this work are all inland on the Antarctic continent in order to avoid the complications of sea emission contamination and melt events. The latter two are at the coast, but experience only a few days each year of temperature above freezing. No other stations have been found that meet the criteria of little or no melting and at least three years of good data records in the period June 2002 - March 2008 (i.e., overlapping with AMSR-E), with the exception of the South Pole station (which cannot be used in this study since orbit geometry prevents most satellites from measuring within disks around the poles). Details are given in Table E.1. The locations are mapped in Fig. E.1.

Stations 89828 is an automated weather station (AWS) located atop Dome C, which is an area of maximum elevation in this region that is known for its stability. Station 89813 is an AWS located 500 km from Dome C, at a lower elevation. Station 89606 is the manned Vostok station and is located 600 km from Dome C. All three of these sites are on the high East Antarctica plateau, a region of little precipitation.

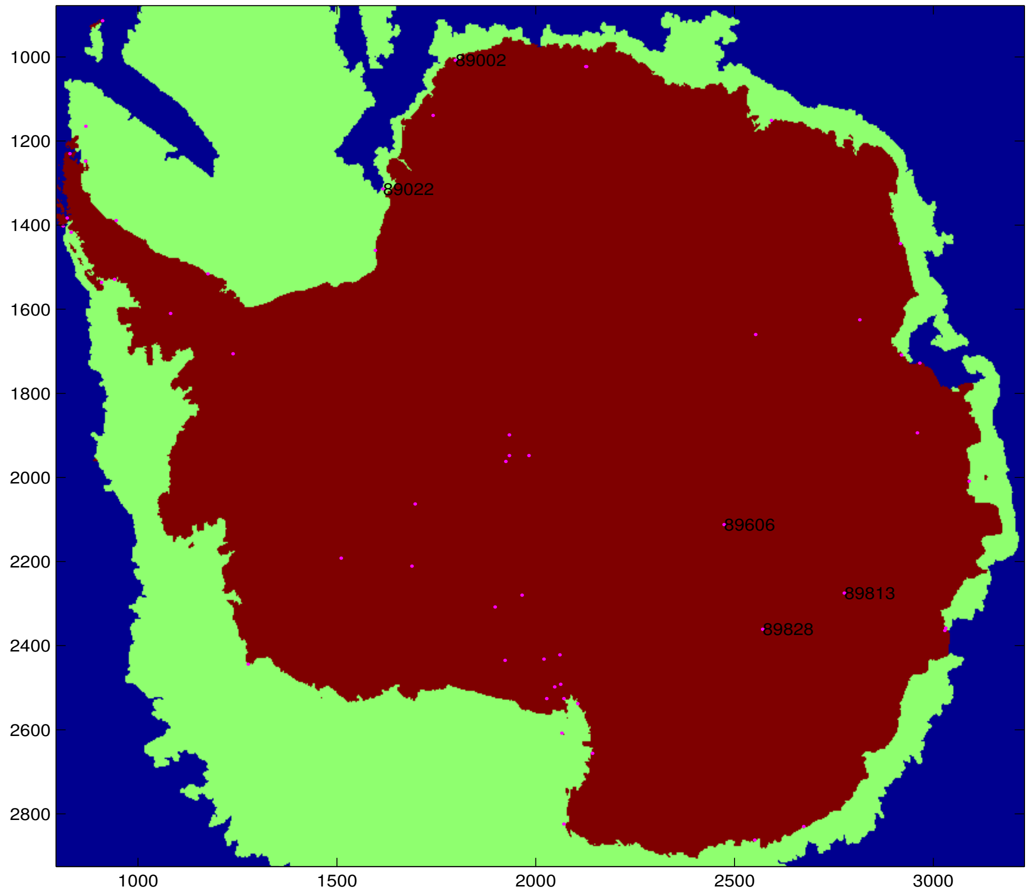

Figure E.1: Map showing locations (pink dots) of weather station sites in Antarctica. Stations labelled with their station number are those that met the necessary criteria for this study. (The green area is roughly the sea ice extent on 1 Jan 2005 and is irrelevant to this dissertation).

Stations 89022 and 89002 are manned and are located near sea level on permanent ice shelves which float on the ocean. The shelf at station 89022 flows seaward at a rate of about 700 m per year. The shelf at station 89002 flow seaward at about 190 m per year, has a thickness of about 200 m, and is almost completely flat. Precipitation is much greater at these stations than at the plateau stations, with station 89022 having an annual snow accumulation of about 1.2 m.

We form daily averages from the station temperature data. Data from stations 89828 and 89813 are required to meet the quality control criteria that at least 12 measurements be available for a day. We also require that the averaged time of day (TOD) of the measurements is not abnormal (specifically, we find the average TOD for the candidate days in a year and then exclude those days on which the average TOD is not within two hours of the group mean).

For Vostok (89606) data, four measurements are usually taken each day, at 6 hour intervals. We require that all four be present. For Neumayer (89002) data, eight measurements are usually taken each day, and we require that all eight be present. For Halley (89022) data, we require that four measurements be present and that the average TOD requirement also be met (measurements there have been collected at least every two hours since March 2005, but less often previous to March 2005).

E.2 Satellite Brightness Temperature Data

The AMSR-E brightness temperature data B ([36], described at http://nsidc.org/data/docs/daac/ae_l2a_tbs.gd.html) can be downloaded from http://www.nsidc.org/data/data_pool/index.html. We use the center latitude and longitude, earth azimuth, and dimensions of each measurement footprint to select only those measurements whose footprint covers the desired ground station latitude and longitude. We manually excise those few data which are obvious outliers (attributed to instrument anomalies). We then form daily averages from these data, with the quality control criteria that the averaged earth azimuth angle for a day is not abnormal (specifically, we find the average angle for all the candidate days and then exclude those days on which the average angle is not within three STD of the group mean).

AMSR-E makes both v-pol and h-pol measurements at frequencies from 7 to 89 GHz. We have chosen to use 37 GHz v-pol measurements because they correlate most highly with surface air temperature. This fact is illustrated for two frequencies in [29] and confirmed by us for AMSR-E 7, 19, 22, and 37 GHz channels. The cause of this higher correlation is probably that 37 GHz measures the *upper* snow and ice better than lower frequencies (i.e., it has a more shallow skin depth) [29] but is not as strongly affected by atmospheric phenomena as 89 GHz radiation.

To prevent the measurements used in this study from including open water or sea ice, we use satellite data covering points located somewhat inland of the two coastal stations. The elliptical footprint of AMSR-E at 37 GHz has dimensions of 14 km by 8 km. Halley station is about 12 km from the ice shelf edge, so we gather satellite measurements centered around a point 0.05 deg (5.5 km) south of the station. Neumayer station is about 13 km south of one ice shelf edge and 4 km west of the other edge, so we gather satellite measurements centered around a point 0.04 deg (4.4 km) south and 0.27 deg (10 km) west of the station.

E.3 Concurrent Data Records

The three years of concurrent T and B records at each site which are used in this work have dates as specified in Table E.1. The number of missing days in each record, usually due to the failure of the T record to meet the quality criteria given above, are also listed. The dates were chosen to minimize the total number of missing days at each station.

Table E.1: Information on the stations and data years used in this work. WMO is the acronym for World Meteorological Organization.

Station name	Dome C II	GC41,Radok	Vostok	Halley Bay	Neumayer
WMO #	89828	89813	89606	89022	89002
Elevation (m)	3250	2761	3488	39	50
Latitude	74.5° S	71.6° S	78.45° S	75.5° S	70.65° S
Longitude	123° E	111.25° E	106.87° E	26.65° W	8.25° W
Year 1 start	9Oct2002	30Oct2002	3Dec2002-13Feb2003	25Feb2005	20Sep2002
Year 1 end	8Oct2003	29Oct2003	14Feb2004 -1Dec2004	24Feb2006	19Sep2003
#missed days	12	3	6	24	22
Year 2 start	3Dec2003	6Nov2003	22Sep2005	25Feb2006	17Nov2003
Year 2 end	1Dec2004	4Nov2004	21Sep2006	24Feb2007	15Nov2004
#missed days	16	7	20	32	23
Year 3 start	15Dec2004	6Nov2004	29Mar2007	25Feb2007	18Feb2005
Year 3 end	14Dec2005	5Nov2005	27Mar2008	24Feb2008	17Feb2006
#missed days	10	6	44	27	11

Bibliography

[1] S. H. Yueh, "Estimates of Faraday rotation with passive microwave polarimetry for microwave remote sensing of earth surfaces," *IEEE Trans. Geosci. Rem. Sens.*, vol. 38, no. 5, pp. 2434–2438, Sep 2000.

[2] J. R. Piepmeier, "Calibration of passive microwave polarimeters that use hybrid coupler-based correlators," *IEEE Trans. Geosci. Rem. Sens.*, vol. 42, no. 2, pp. 391–400, Feb 2004.

[3] A. Tarantola, *Inverse problem theory and methods for model parameter estimation.* Philadelphia, PA: SIAM, 2005, p. 44-49, 173, and 179-180. The book is conditionally available at http://www.ipgp.jussieu.fr/%7Etarantola/Files/Professional/SIAM/index.html.

[4] D. L. Hudson, J. R. Piepmeier, and D. G. Long, "Polarization rotation correction in radiometry: an error analysis," *IEEE Trans. Geosci. Rem. Sens.*, vol. 45, no. 10, pp. 3212–3223, Oct 2007.

[5] D. L. Hudson and D. G. Long, "Optimal estimation of calibration parameters in polarimetric microwave radiometers," *IEEE Trans. Geosci. Rem. Sens.*, vol. 46, no. 10, pp. 3223–3237, Oct 2008.

[6] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive.* Norwood, MA: Artech House, 1981, vol. 1.

[7] ——, *Microwave Remote Sensing: Active and Passive.* Norwood, MA: Artech House, 1986, vol. 2.

[8] A. Papoulis, *Probability, Random Variables, and Stochastic Processes, Fourth ed.* New York, NY: McGraw-Hill, 2002.

[9] E. T. Jaynes, *Probability Theory: The Logic of Science.* Cambridge, U.K.: Cambridge Univ. Press, 2003.

[10] D. M. Le Vine and S. Abraham, “The effect of the ionosphere on remote sensing of sea surface salinity from space: absorbtion and emission at L-band,” *IEEE Trans. Geosci. Rem. Sens.*, vol. 40, no. 4, pp. 771–782, Apr 2002.

[11] T. Meissner and F. J. Wentz, “Polarization rotation and the third Stokes parameter: the effects of spacecraft attitude and Faraday rotation,” *IEEE Trans. Geosci. Rem. Sens.*, vol. 44, no. 3, pp. 506–515, Mar 2006.

[12] A. J. Gasiewski and D. B. Kunkee, “Calibration and applications of polarization-correlating radiometers,” *IEEE Trans. Microw. Theory Tech.*, vol. 41, no. 5, pp. 767–773, May 1993.

[13] I. Corbella, F. Torres, A. Camps, A. Colliander, M. Martín-Neira, S. Ribó, K. Rautiainen, N. Duffo, and M. Vall-llossera, “MIRAS end-to-end calibration: application to SMOS L1 processor,” *IEEE Trans. Geosci. Rem. Sens.*, vol. 43, no. 5, pp. 1126–1134, May 2005.

[14] A. Team, *Aquarius Selected Instrument Concept.* Greenbelt, MD: Goddard Space Flight Center, 2005, [Online]. Available: http://aquarius.gsfc.nasa.gov/pdf/instrument.pdf.

[15] S. Chandrasekhar, *Radiative Transfer*, ser. The International series of monographs on physics. Oxford, England: Oxford, Claredon Press, 1950.

[16] S. Ribó and M. Martín-Neira, “Faraday rotation correction in the polarimetric mode of MIRAS,” *IEEE Trans. Geosci. Rem. Sens.*, vol. 42, no. 7, pp. 1405–1410, Jul 2004.

[17] Y. Kerr, P. Waldteufel, and F. Cabot, “SMOS geolocation using natural targets,” Sep 2004, [Online]. Available: http://www.cesbio.ups-tlse.fr/data_all/SMOS-doc/geolocation.pdf.

[18] D. Entekhabi, E. G. Njoku, P. Houser, M. Spencer, T. Doiron, Y. Jim, J. Smith, R. Girard, S. Belair, W. Crow, T. J. Jackson, Y. H. Kerr, J. S. Kimball, R. Koster, K. C. McDonald, P. E. O'Neill, T. Pultz, S. W. Running, J. Shi, E. Wood, and J. van Zyl, "The hydrosphere state (Hydros) satellite mission: an earth system pathfinder for global mapping of soil moisture and land freeze/thaw," *IEEE Trans. Geosci. Rem. Sens.*, vol. 42, no. 10, pp. 2184–2195, Oct 2004.

[19] H. Weil, "The distribution of radial error," *The Annals of Mathematical Statistics*, vol. 25, no. 1, pp. 168–170, 1954.

[20] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991, exercise 6-16.

[21] K. S. Miller, *Multidimensional Gaussian Distributions*, ser. SIAM Series in Applied Mathematics. New York: Wiley, 1964.

[22] S. O. Rice, "Mathematical analysis of random noise – conclusion," *Bell Systems Tech. J.*, vol. 24, no. 1, pp. 46–156, 1945.

[23] S. H. Yueh, W. J. Wilson, S. J. Dinardo, and S. V. Hsiao, "Polarimetric microwave wind radiometer model function and retrieval testing for WindSat," *IEEE Trans. Geosci. Rem. Sens.*, vol. 44, no. 3, pp. 584–596, Mar 2006.

[24] S. S. Sobjaerg, J. Rotboll, and N. Skou, "Measurement of wind signatures on the sea surface using an L-band polarimetric radiometer," in *Proc. IEEE Int. Geoscience and Remote Sensing Sympos.*, vol. 3, 2002, pp. 1364–1366.

[25] J. R. Piepmeier, private communication, 2006.

[26] T. K. Moon and W. C. Stirling, *Mathematical methods and algorithms for signal processing.* Upper Saddle River, New Jersey: Prentice Hall, 2000, 2nd printing, Theorem 6.2 and Theorem 12.8.

[27] B. D. Flury, "Acceptance-rejection sampling made easy," *SIAM Review*, vol. 32, no. 3, pp. 474–476, Sep 1990.

[28] S. Surdyk, “Using microwave brightness temperature to detect short-term surface air temperature changes in Antarctica: An analytical approach,” *Remote Sens. Environ.*, vol. 80, no. 2, pp. 256–271, 2002.

[29] C. A. Shuman, R. B. Alley, S. Anandakrishnan, and C. R. Stearns, “An empirical technique for estimating near-surface air temperature trends in central greenland from ssm/i brightness temperatures,” *Rem. Sens. Env.*, vol. 51, no. 2, pp. 245–252, Feb 1995.

[30] J. C. Comiso, “Variability and trends in antarctic surface temperatures from in situ and satellite infrared measurements,” *J. Climate*, vol. 13, pp. 1674–1696, 15 May 2000.

[31] H. Zwally, “Microwave emissivity and accumulation rate of polar firn,” *J. Glaciol.*, vol. 18, pp. 195–215, 1977.

[32] D. G. Long and M. R. Drinkwater, “Azimuth variation in microwave scatterometer and radiometer data over antarctica,” *IEEE Trans. Geosci. Rem. Sens.*, vol. 38, no. 4, pp. 1857–1870, Jul 2000.

[33] D. Hudson, J. Piepmeier, and D. Long, “Polarization rotation correction in radiometry: An extended error analysis,” in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, vol. 7, Aug 2006, pp. 2305–2308.

[34] SMAP Team, *Soil Moisture Active Passive.* Pasadena, CA: Jet Propulsion Laboratory, 2009, [Online]. Available: http://smap.jpl.nasa.gov/.

[35] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*, 4th ed. New York: Academic Press, 1965, sec. 3.462.

[36] P. Ashcroft and F. Wentz, *AMSR-E/Aqua L2A Global Swath Spatially-Resampled Brightness Temperatures (Tb) V08, June 2002 - Feb 2008.* Boulder, CO, USA: National Snow and Ice Data Center, 2008, updated daily, digital media.