Design and analysis of radiometric instruments using high-level numerical models and genetic algorithms

by

Ira Sorensen, M.S.

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Mechanical Engineering

Dr. J. R. Mahan
Dr. Kory Priestley
Dr. Elaine Scott
Dr. Brian Vick
Dr. Zafer Gurdal

December 2002

Blacksburg, Virginia

Keywords: Radiometry, Remote Sensing, Genetic Algorithms, Monte Carlo Ray Trace

Design and analysis of radiometric instruments using high-level numerical models and genetic algorithms

Ira Sorensen, Ph.D.
Virginia Polytechnic Institute and State University, 2002
Advisor: J. R. Mahan

Abstract

A primary objective of the effort reported here is to develop a radiometric instrument modeling environment to provide complete end-to-end numerical models of radiometric instruments, integrating the optical, electro-thermal, and electronic systems. The modeling environment consists of a Monte Carlo ray-trace (MCRT) model of the optical system coupled to a transient, three-dimensional finite-difference electrothermal model of the detector assembly with an analytic model of the signal-conditioning circuitry. The environment provides a complete simulation of the dynamic optical and electrothermal behavior of the instrument. The modeling environment is used to create an end-to-end model of the CERES scanning radiometer, and its performance is compared to the performance of an operational CERES total channel as a benchmark. A further objective of this effort is to formulate an efficient design environment for radiometric instruments. To this end, the modeling environment is then combined with evolutionary search algorithms known as genetic algorithms (GA's) to develop a methodology for optimal instrument design using high-level radiometric instrument models. GA's are applied to the design of the optical system and detector system separately and to both as an aggregate function with positive results.

Acknowledgments

I would first like to thank my academic advisor, Dr. J.R. Mahan, for his help and patience throughout my graduate studies. Under his guidance I have been allowed to grow both as a researcher and as an individual.

I would like to thank Dr. Elaine Scott, Dr. Brian Vick and Dr. Zafer Gurdal for their willingness serve on my advisory committee. I would also like to extend my thanks to Dr. Kory J. Priestley of the NASA Langley Research Center for serving on my committee as well as serving as my mentor in the NASA Langley Research Center Co-op Program, and for providing a challenging and stimulating environment during my stay there. I would also like to thank the NASA Langley Research Center, which supported this work under grant number NAG-1-2128.

I would like to thank my colleagues, past and present, in Thermal Radiation Group, both at Virginia Tech and at the NASA Langley Research Center: Félix Nevárez, Amie Smith, Dwight Smith, Bernardo Carnicero, Jose Dobarco, Maria Santa Maria, Mamadou Barry, Stéphanie Weckmann, Katherine Coffey, Edwin Ayala, Joel Barreto, and Martial Haeffelin. Their generous help and friendship was invaluable in assisting my research.

I would like to thank my parents, Duane and Joyce, as well as my brothers, Daniel and Jacob for their confidence and support throughout the years. Finally, I would like to thank my wife, Cristina, without whose love and support the realization of my goals and aspirations would not have been possible.

Ira Sorensen

Virginia Polytechnic Institute and State University

December 2002

Contents

List of Figures

List of Tables

Nomenclature

Symbols:

A	Area (m^2)
B	Thermistor material constant (K)
B^*	Normalized inverse of the blur circle diameter (-)
C	Capacitance (μF)
CI	Confidence interval (-)
D	Detectivity (W^{-1})
D^*	Normalized detectivity (cm $Hz^{1/2}$/W)
D'_{ij}	Distribution factor between surface elements i and j (-)
E	Electric potential (V)
E_{rel}	Relative error (-)
f	Frequency (Hz)
FI	Fidelity interval (-)
G	Instrument Gain ($W/m^2 \cdot$ sr $\cdot$ counts)
I	Electric current (A)
I	Radiance ($W/m^2 \cdot$ sr)
J	Current Density (A/m^2)
k	Thermal conductivity ($W/m \cdot K$)
l	Scene resolution (m)
m	Digital counts (-)

O^*	Normalized optical throughput (-)
Obj	Objective function
P	Temperature coefficient of resistance (K^{-1})
q	Heat flow rate (W)
$\dot{q}$	Volumetric heat generation (W/m^3)
r	Distance between scene and instrument aperture (m)
r_s	Specularity ratio (-)
R_e	Resistance (Ω)
R	Responsivity (V/W)
s	Standard deviation (-)
S	Spectral response function (-)
S_{AB}	Seebeck coefficient of materials A and B (V/K)
t	Time (s)
T	Temperature (K)
v	Field of view (rad)
V	Voltage (V)
Y	Instrument output (V)

Greek:

α	Absorptivity (-)
δ	Kronecker delta function (-)
λ	Wavelength (μm)
Ω	Solid angle (sr)
ρ	Reflectivity (-)
ρ_e	Electrical resistivity ($\Omega\cdot$ m)
ρC_p	Heat capacity per unit volume ($J/m^3\cdot$ K)
σ	Stefan-Boltzmann coefficient ($W/m^2\cdot K^4$)

σ_e	Electrical conductivity $(\Omega \cdot \mathrm{m})^{-1}$
τ_{th}	Time constant (s)
τ	Transmissivity (-)

Subscripts:

bb	Blackbody
f	Filtered
i	Surface element i
j	Surface element j
n	Noise
$p-p$	Peak-to-peak
rms	Root-mean-square
T	Thermal
x	x-direction
y	y-direction
z	z-direction

Chapter 1

Introduction

Increased industrialization over the last century, while bringing a myriad of benefits to society, has greatly compromised the world's social structure and its ecosystem. Pollution of our rivers, oceans, soil, and atmosphere has increased significantly during this century.

One area of increasing concern is the ongoing pollution of the atmosphere. Continued emissions of CFC's, as well as the burning of fossil fuels, increases the level of CO_2 in the atmosphere, which in turn has led to an increased concern over global warming due to the "greenhouse effect."

In order to understand the effects of these changes in our climate, we must first seek to understand the climate system itself. For example, it has been hypothesized that the greenhouse effect causes, through warming of the lower troposphere, increased concentration of water vapor in the atmosphere. Because water is itself a greenhouse gas, this raises the possibility of "runaway" global heating. However, increased water vapor concentration should increase the production of clouds in the atmosphere, which could then have an overall cooling effect on the climate. Which of these two trends might actually dominate is a matter of ongoing debate [Kandel, 1998].

In an effort to better understand the climate system, the international scientific community

has initiated a series of programs aimed at increasing our knowledge and understanding of the climate through observation of climatic changes and radiation in the atmosphere.

As part of the National Aeronautics and Space Administration's (NASA) ongoing efforts to monitor the earth radiation budget, the Clouds and the Earth's Radiant Energy System (CERES) instrument [Wielicki et al., 1996; Carman et al., 1992] provides broadband measurements of reflected solar and earth-emitted longwave radiances. These measurements are a continuation of the Earth Radiation Budget Experiment (ERBE) measurements [Barkstrom and Hall, 1982; Barkstrom and Smith, 1986] and are used in scientific investigations to determine top-of-the-atmosphere (TOA) heat fluxes, radiative forcing properties of clouds, Earth surface radiation budget, and divergence throughout the atmosphere.

Due to the importance of these measurements to the scientific community, it is neccessary to obtain a thorough understanding of the behavior of the instrument and the underlying physics governing that behavior. This can be achieved using detailed analytical models based on first principles of physics. High-level numerical models of the instrument, modeling its optical, thermal, and electrical behavior, can be used to gain an increased understanding of its performance to aid in both the design process and the analysis of operational data. This results in improved instrument concepts with operational characteristics that maximize the accuracy of the instrument's measurements.

1.1 Radiometric instrument modeling

The Thermal Radiation Group (TRG), under the guidance of Dr. J.R. Mahan of the Department of Mechanical Engineering at Virginia Polytechnic Institute and State University, has been involved in the development of optical, radiative, thermal conduction and electronic models of spaceborne radiometers for the last twenty-five years. High-level dynamic electrothermal modeling efforts include NASA's ERBE and CERES instruments.

Early efforts to study radiometric instruments include Fanney [1975], who peformed an ex-

perimental study of the optical and thermal characteristics and performance of the spherical detectors proposed for the LZEEBE (Long-Term, Zonal Earth Energy Budget Experiment) instrument. A detailed thermal model of the spherical detectors was developed by Rasnic [1975], and later upgraded by Passwaters [1976] to incorporate realistic radiative exchange between the Earth's radiative field and the detector surface.

Subsequent models of the optical systems of radiometric instruments have typically been developed using the Monte Carlo ray-trace method (MCRT). This is a statistical ray-trace method that uses probabilities to represent the radiative properties of surfaces and to calculate distibution factors, which are similar but superior to configuration factors in that they capture the optical properties as well as the geometry of an enclosure. Eskin [1981] was the first to apply the MCRT method to model a cavity radiometer, and he also developed the idea of the distribution factor with Mahan [Mahan and Eskin, 1984]. Kowsary [1989] peformed a study of the radiative characteristics of a diffuse-specular spherical cavity using an exact analytic method and verified his reults using the MCRT method. Walkup [1996] used the MCRT method to develop a virtual optical workbench to aid in the design and analysis of radiometric imaging systems.

Several students devloped radiometric models of the Earth Radiation Budget Experiment (ERBE) scanning and non-scanning radiometers. Tira [1987] developed a distribution factor-based model of the ERBE optical system and a finite-element model of the ERBE detector, and he also performed a dynamic study of the instrument calibration. In his doctoral work, Tira [1991] developed improved models of the ERBE staring channels to simulate their solar calibration and dynamic response during the pitch-over maneuver. He also developed a model of the staring active cavity radiometer (ACR) to study its thermal noise and transfer function. Meekins [1990] performed a detailed analysis of the ERBE scanning radiometer optical system and its radiative characteristics using the MCRT method. Haeffelin [1993] developed a high-level electrothermal model of the ERBE thermistor bolometer detectors to study the dynamic thermal and electrical behavior of the instrument.

The Clouds and Earth's Radiant Energy System (CERES) instrument provided fruitful research for a new generation of students. Bongiovi [1993] developed a Monte Carlo ray-race model of the CERES optical system and performed an analytical study of its radiative behavior and performance. Savransky [1996] developed a thermal model of the instrument housing and structure using the finite-element method to predict transient temperature response and the effect of thermal noise on the performance of the instrument. Haeffelin [1997] developed a three-dimensional dynamic electro-thermal diffusion model of the CERES thermistor bolometer. He used results from Bongiovi's optical model as an input to his dynamic electrothermal model to perform detailed studies of the instrument's electrical and thermal performance. Priestley [1997] extended the CERES electrothermal detector model by incorporating the reference detector and intervening substrate, and improved its performance by updating model parameters based on available calibration data. He also integrated the model with a short-wavelength atmospheric Monte Carlo ray-trace model, developed by Villeneuve [1996], which provides a simulation of the earth-reflected radiation field at the top of the atmosphere, to perform realistic instrument simulations and provide an accurate assessment of the instrument's in-flight performance. Coffey [1998] developed an MCRT model of the CERES optical system and studied the instrument's predicted performance when the spherical mirrors were replaced with hyperbolic mirrors and the single precision aperture was replaced with multiple precision apertures.

In addition to developing detailed models of existing instruments, several members of the TRG have been involved in using numerical models to develop new instrument concepts. Weckmann [1997] developed a detailed finite-element thermal model of a new thermopile detector concept. Sánchez [1998] developed an MCRT-based optical model of a wedge-shaped cavity radiometer, and Sorensen [1998] integrated the thermal and optical models into a complete end-to-end model of a wedge-shaped cavity radiometer and developed a numerical algorithm to remove the optical crosstalk due to diffuse scattering in the cavity. Mahan et al. [1999] developed an MCRT-based optical model of a novel aureolimeter instrument concept and performed a detailed optical analysis. Barry [1999] studied thermal conduction issues

in thin-film thermal detectors, and Dobarco-Otero [2000] studied absorption and thermal diffusion issues of long-wavelength radiation in thin-film detectors. Detailed analytic models of pyronometers used to provide surface radiance measurements were developed by Carnicero [2001] and Smith [1999] in addition to experimental studies of the instruments' thermal behavior.

1.2 Goals and motivation

Over the last five years, the TRG has been involved in an effort to develop numerical models to test new instrument concepts for use in the next generation of spaceborne instruments, including possible use on the next generation of CERES applications. The technology development has consisted of two major efforts:

- development of interactive computer environments to aid in the development of optical and electrothermal end-to-end numerical instrument models to predict the behavior and performance of remote sensing instruments, and
- creation of a computer-based optimization environment that utilizes these models to develop optimal instrument concepts and models under a variety of goals and constraints.

Toward these ends, a Monte-Carlo ray trace (MCRT) design environment was developed by Nevárez [2001] and a statistical methodology for determining the accuracy of high-level end-to-end instrument models is currently being developed [Sánchez et al., 1999, 2000, 2001; Sánchez, 2002]. Efforts to incorporate bi-directional surface properties and diffraction effects into the design environment are also underway.

The aims of the research presented in this dissertation are to extend the capabilities of the current generation of high-level numerical instrument models by achieving the following:

- develop a generic instrument modeling environment that integrates the optical and electrothermal system models and allows the user to test different configurations and components in the optical and detector systems,
- use the environment to benchmark a complete end-to-end model of the CERES instrument against measured performance parameters from an operational CERES instrument, and
- integrate the modeling environment with intelligent search algorithms (genetic algorithms) to provide an instrument design and optimization tool to aid in the development of new instrument concepts.

1.3 Radiometric instrument modeling environment

The goal of the effort presented here is to develop a radiometric instrument modeling and design environment. This environment will allow the user to develop end-to-end models of radiometric instruments, integrating models of the optical, thermal, and electrical systems of the instrument. The user can select from five mirror configurations within a classical two-mirror system:

1. a Cassegrain configuration featuring a parabolic primary mirror and a hyperbolic secondary mirror
2. a Gregorian configuration featuring a parabolic primary mirror and an ellipsoid secondary mirror
3. a Dall-Kirkham configuration featuring an ellipsoid primary mirror and a spherical secondary mirror
4. a Ritchey-Chrétian configuration featuring two hyperbolic mirrors

5. a modified Cassegrain consisting of two spherical mirrors (CERES configuration)

The user provides the mirror dimensions and locations within a telescope system based on the CERES instrument design, and also defines the optical properties of the system, such as absorptivity and specularity, for both gray and spectral models. The optical model then uses the Monte Carlo ray-trace method to simulate the optical performance of the instrument and determine the distribution of incident radiation throughout the instrument.

The current modeling environment limits the user to a thermistor bolometer detector system, modeled on the detectors used on the current generation of CERES instruments. The detector assembly consists of two thermistor bolometers separated by an aluminum substrate, with one detector acting as an active detector and the other acting as a reference detector. Each detector consists of a thin thermistor layer separated from a substrate by a thin sheet of Kapton, which acts as a thermal impedance, and a thin layer of highly absorptive black paint above the thermistor, which absorbs incident radiation. The detector is described in more detail in Chapter 4. The user can define all thermophysical and electrical properties of the system, as well as layer thicknesses and all spatial dimensions. A three-dimensional transient finite-difference thermal model is then created to determine the temperature distribution at each time-step. Two separate finite-difference electric models are created to represent the active and reference thermistor layers. These are two-dimensional steady-state models, whose state is calculated at each discrete time step. The electric models are coupled to the thermal model, and at each time step the temperature-dependent resistivity of each element in the electric models are updated and the electric field in each is recalculated. The active and reference thermistors are in adjacent arms of a Wheatstone bridge circuit, and the results of the electric models are used in models of the bridge circuit, pre-amplifier, and analog-to-digital converter to calculate the instrument output at each time step.

The environment can provide the values of various parameters representing the performance of the instrument, including the spatial distribution of energy on the detector flake, absorbed energy on various surfaces, the optical point spread function, the instrument point spread

function, the temperature distibution throughout the detector assembly, the electric field in both thermistors, and the signal output of the instrument. The terminology is defined at the appropriate location elsewhere in this thesis.

1.4 Organization of dissertation

The dissertation is divided into seven chapters. The first chapter provides an overview of the motivation for and the aims of the current research. The second chapter provides an overview of radiometry and radiometric instruments, including relevant theory and published works pertaining to the modeling and design of the optical systems and detectors used in these instruments. The third chapter provides a brief outline of the development and mechanisms of genetic algorithms (GA's), which are intelligent search algorithms based on evolutionary principles, that have been adapted for use in radiometric design in the current effort. The fourth chapter provides an in-depth description of the modeling environment, detailing the development of both the optical and electrothermal models. The fifth chapter describes the application of the modeling environment to the Clouds and Earth's Radiant Energy System (CERES) instrument, and comparisons of the modeled performance to the actual instrument performance. The sixth chapter discusses the application of genetic algorithms to the modeling environment to provide a tool for instrument design. An analysis of the GA's performance for different test cases and objectives is presented, as well as an analysis of hybrid genetic algorithms and their performance. Finally, chapter seven provides conclusions and recommendations issuing from the current work.

Chapter 2

Radiometric Instruments

This chapter provides the reader with a brief overview of radiometric instruments and radiomety. Topics include general properties of optical systems and detectors, including a review of optical theory and various mirror configurations used in radiometric instruments. The different types of thermal detectors and associated performance parameters are discussed. This chapter is not meant to be an exhaustive discussion of these topics, but rather is an introduction to topics treated in more detail in later chapters.

2.1 General properties of radiometric instruments

Radiometric instruments consist of an optical system to gather incident radiation, a detector system to convert the absorbed energy into an electric signal, and an electronic signal conditioning circuit for processing the signal. The optical system is typically either some type of telescope configuration that focuses incident radiation onto a single detector or array of detectors located in the focal plane, or a cavity whose shape and surface optical properties combine to trap incident radiation, which is then ultimately absorbed on the detector after a series of reflections. The detector system typically consists of thermal detectors that absorb

incident radiant power and produce a proportional electric signal. In this implementation, the primary thermal detectors used include thermopile detectors, which produce a signal using the thermoelectric effect; bolometers, which use temperature-dependent resistivity to produce a signal; and pyroelectric detectors, which use the change in electrical polarization of a crystal due to a change in its temperature to produce an electric signal.

Various figures of merit are used to evaluate the performance of radiometric instruments. The performance of an optical system can be characterized by its the optical throughput, its field-of-view, its spectral response function, and the size of its blur circle, to name but a few. Detector performance evaluators include responsivity, time response, noise-equivalent power (NEP), and signal-to-noise ratio.

2.2 Optical systems

The primary goal of the optical system is to gather incident radiation and direct it to the detector system. The optical system consists of baffles, mirrors, lenses, filters, and apertures assembled in a telescope configuration. The optical system defines the spatial and directional characteristics of the instrument and, along with the detector absorber layer, strongly influences the spectral characteristics as well. Considerations when designing an optical system for radiometric instruments include:

- the system should be designed to provide a specified field-of-view (FOV); the FOV can typically be defined by a precision aperture aligned normal to the optical axis near the detector plane,
- the system should strive to maximize the optical throughput within the prescribed field-of-view and desired wavelength interval,
- the blur circle at the detector plane should generally be minimized (although it is sometimes purposefully defocused to reduce high-frequency aliasing), and

- the sytem should be as spectrally flat as possible within the desired wavelength interval. By "spectrally flat" it is meant that the instrument response is independent of wavelength in a specified interval.

The optical systems considered in this work consist only of two-mirror telescopes. This is justified because two-mirror telescopes are a popular choice in radiometric instrumentation applications, and the scope of the present effort needs to be limited. These systems consist of a primary mirror, which reflects incident radiation to a secondary mirror located in front of the primary mirror, which in turn reflects the radiation to a detector typically located behind the primary mirror. Five two-mirror systems are considered in the current effort:

- a Cassegrain telescope, consisting of a concave parabolic primary mirror and a convex hyperbolic secondary mirror,
- a Gregorian telescope, consisting of a concave parabolic primary mirror and a concave elliptical secondary mirror,
- a Dall-Kirkham telescope, consisting of a concave elliptical primary mirror and a convex spherical secondary mirror,
- a Ritchey-Chrétian telescope, consisting of a concave hyperbolic mirror and a convex hyperbolic mirror, and
- a CERES-like telescope, consisting of a concave spherical primary mirror and a convex spherical seconday mirror.

A typical two-mirror Cassegrain-type system is illustrated in Figure 2.1. The Cassegrain system is commonly used in radiometric systems because of its ability to sharply reduce spherical aberrations. The parabolic primary mirror eliminates third-order aberrations from the primary image, and the hyperbolic element has the property that all rays aimed at its primary focus will be reflected to its secondary focal point without aberration on-axis

[Wolfe, 1996]. However, the Cassegrain telescope is costly to fabricate and is often replaced with a two-mirror system, known as the Dall-Kirkham configuration, having a spherical secondary mirror. This system has a reduced performance compared to a true Cassegrain telescope, but is less costly to produce because of the relative ease of forming a spherical mirror. The Ritchey-Chrétian configuration takes the alternate tack, using two hyperbolic mirrors to provide superior performance to a Cassegrain, but with a significant increase in cost. Another variation is to use a system in which the primary or both mirrors are spherical, as this is even less costly to produce than the Dall-Kirkham. However, this option comes with another reduction in performance. The use of two spherical mirrors results in a reduction in off-axis aberrations, which results in a more sharply defined field-of-view, but with an increase in on-axis aberrations. However, this is tolerable if the instrument is not an imaging instrument. The Gregorian configuration is easier to manufacture due to having two convex mirrors; however, it is typically longer than the other systems considered and so is not used as frequently in space-based systems.

2.3 Optical system parameters

This section provides a brief explanation of various parameters and figures of merit used to evaluate the performance of optical systems. These are used later in an objective function, in concert with a search algorithm, to provide an instrument design and optimization tool. These figures of merit include the blur circle diameter at the detector plane, the instrument field-of-view (FOV), and the optical point spread function (OPSF).

2.3.1 Blur circle

The quality of an optical system can be expressed by the size of the image formed by a point source. The image formed is not a point but has a finite size due to diffraction effects and spherical aberrations. The periphery of the image defines what is referred to as the

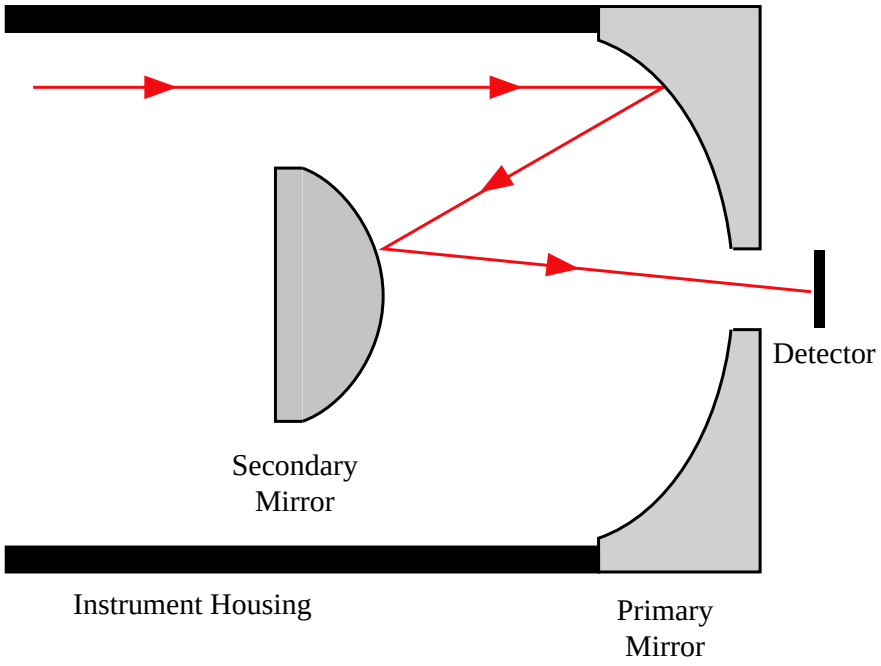

Figure 2.1: Illustration of a typical Cassegrain-type two-mirror system

blur circle, and its dimensions can be expressed in angular units or in terms of its linear diameter [Wyatt, 1987]. An example of a blur circle for the CERES instrument produced by on-axis collimated radiation is shown in Figure 2.2. This image was created from a ray-trace model of the CERES instrument, and the crosses represent the intersection of individual rays entering the instrument aperture with the focal plane, after being traced through the optical system. It should be noted that the ray-trace model does not include diffraction effects. The three radial sectors with a lesser concentration of crosses represent the image formed by the three legs of the "spider," the familiar name given to the structure which supports the secondary mirror. Note that the diameter of the blur circle in this case is approximately 0.07 mm, which is significantly less than 1.6 mm, the length and width of the CERES detector "flake".

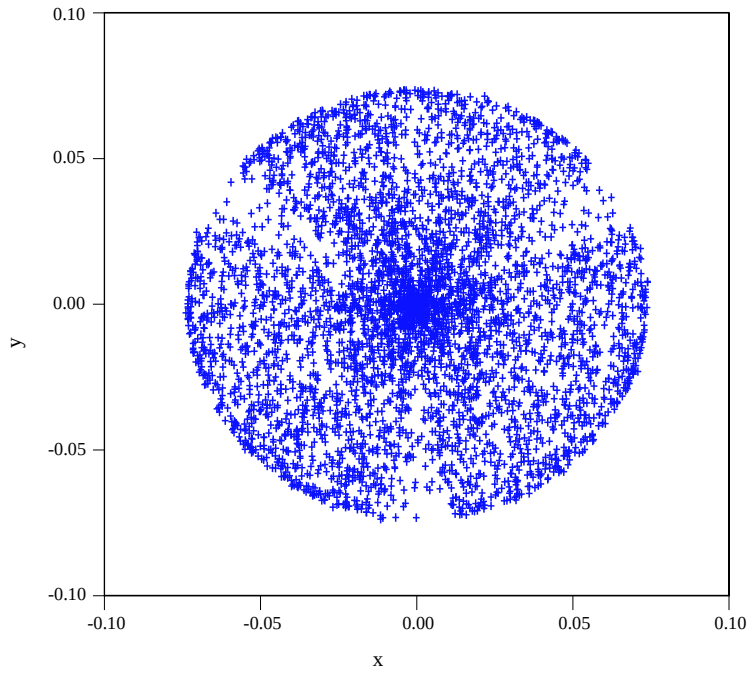

Figure 2.2: Blur circle for on-axis collimated radiation in a two-spherical-mirror system

2.3.2 Field-of-view and optical point spread function

The spatial response of the instrument is defined by the field-of-view (FOV). The field-of-view is the angular resolution of the instrument. The differential field-of-view, dv, is defined as the ratio of the differential scene resolution, dl, to the distance from the instrument aperture to the scene, r [Chen, 1997], that is

$$dv = \frac{dl}{r} \tag{2.1}$$

The instantaneous field-of-view, v, is related to the solid angle, Ω, by [Chen, 1985]

$$\Omega = v^2 \ (sr) \tag{2.2}$$

The field-of-view must be matched to the source of incident radiation, and is typically specified prior to the instrument design. A "scanning" instrument is gimbaled so that the point

where its optical axis intersects the scene sweeps across the scene. The optical point spread function (OPSF) is related to the field-of-view, and is defined as the angular response of the instrument to collimated radiation. The OPSF, which is an inherent property of the optical system, differs from the field-of-view in that the FOV is expressed as a value while the OPSF is expressed as a distribution along the scan and cross-scan angles. The scan angle is an angle measured from nadir along the instrument's scan direction, and the cross-scan angle is measured from nadir in the direction perpendicular to the scan direction. These angles are illustrated in Figure 2.3. An example of the OPSF for the CERES optical system is shown in Figure 2.4. The figure shows the normalized response of the instrument to collimated radiation coming from various scan and cross-scan angles, and was created from the MCRT model of the CERES instrument described elsewhere in this dissertation. The instrument's response is uniform over a range of ± 0.8 deg in the scan direction and ± 1.6 deg in the cross-scan direction. The shape of the OPSF closely matches the shape of the diamond-shaped precision aperture used on CERES.

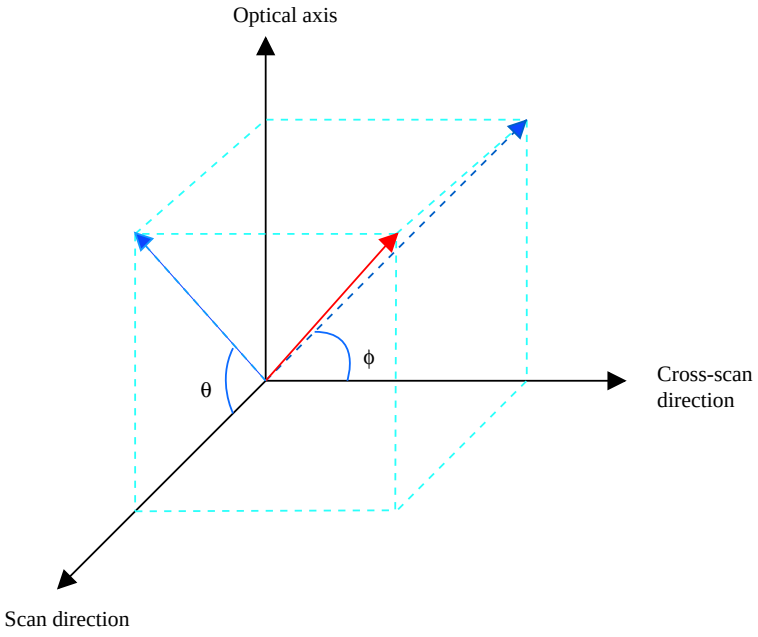

Figure 2.3: Illustration of scan, θ, and cross-scan, ϕ,angles

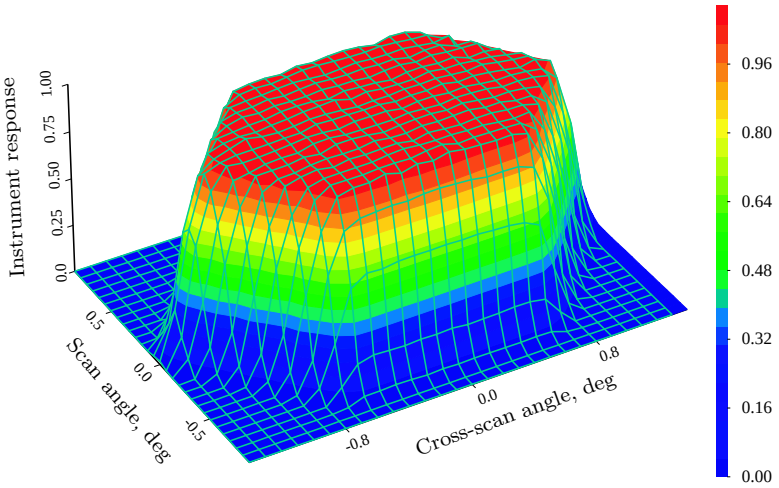

Figure 2.4: Optical point spread function (OPSF) surface plot for the CERES instrument

2.4 Thermal detectors

The input signal, or flux, to a filtered radiation detector is

$$Q(\lambda) = \alpha_\lambda(\lambda)\,\Phi_\lambda(\lambda) \qquad (\mathrm{W/m^2}) \tag{2.3}$$

where $\Phi_\lambda(\lambda)$ is the incident monochromatic radiant flux ($\mathrm{W/m^2 \cdot sr}$) and $\alpha_\lambda(\lambda)$ is the monochromatic absorptivity (-) of the detector surface at a particular wavelength λ.

The absorptivity of the detector can be enhanced by two means. The first is through the use of a "black" coating, i.e., one with an absorptivity approaching unity, on the detector surface. The second is through the use of a cavity, which causes the incident flux to strike the detector surface multiple times through mostly specular (mirror-like) reflections.

Two major categories of radiation detectors, thermal detectors and photon detectors, may

be identified. Thermal detectors transform the absorbed radiant energy into sensible heat, producing a temperature rise that can be directly converted into an electrical signal. In photon detectors, absorbed photons excite electrons in the detector material, thereby increasing their energy level. If the energy of an individual electron is raised sufficiently, it becomes a free electron that may be extracted from the detector, thereby producing a measurable photoelectric current.

The *responsivity*, R, of the detector, as used in the current work, is defined as the signal output, Y (V), per unit of absorbed radiation, Q (W), and is represented by [Lenoble, 1993]

$$R = \frac{Y}{Q} \qquad \text{(V/W)} \tag{2.4}$$

The output of the detector always contains a certain amount of noise. A portion of this noise arises from the measurement procedure and can be minimized. The remainder is noise that is inherent to the physical properties of the detector, and is known as *fundamental noise*. The output signal of the detector can be expressed as the mean value $\langle Y \rangle$ defined

$$\langle Y \rangle = \frac{1}{n} \sum_{i=1}^{n} Y_i \tag{2.5}$$

and the noise is expressed as the rms noise $(\langle \Delta Y^2 \rangle)^{1/2}$ defined

$$N = (\langle \Delta Y^2 \rangle)^{1/2} = \left[\frac{1}{n} \sum_{i=1}^{n} (Y_i - \langle Y \rangle)^2 \right]^{1/2} \tag{2.6}$$

The *signal-to-noise ratio* is defined as the ratio of the mean signal to the rms noise,

$$SNR = \frac{\langle Y \rangle}{(\langle \Delta Y^2 \rangle)^{1/2}} \tag{2.7}$$

The *noise equivalent power* (NEP) is the detector input that produces a $SNR = 1$, and the *detectivity* is given by

$$D = \frac{1}{NEP} \tag{2.8}$$

Detector responsivity is typically wavelength dependent, but different detectors can be characterized using the *normalized detectivity*

$$D^* = D\,(A\,\Delta f)^{1/2} \tag{2.9}$$

with A being the sensitive area of the detector and Δf the measurement bandwidth.

The output signal, Y, of a thermal detector is induced by a temperature variation ΔT caused by the absorbed power. The responsivity of the thermal detector can be expressed as [Lenoble, 1993]

$$R = \left(\frac{\Delta T}{Q}\right)\left(\frac{Y}{\Delta T}\right) \qquad (V/W) \tag{2.10}$$

The factor $\Delta T\,/\,Q$ is a physical property that depends only on the thermal characteristics of the detector.

If the temperature of the detector is raised by ΔT_o with respect to a reference and then allowed to decay without any external heat source present, the exponential temperature decay is given by

$$\Delta T(t) = \Delta T_o\,\exp\left[-\frac{t}{\tau}\right] \tag{2.11}$$

where $\tau = C_T\,R_T$ is the thermal time constant of the detector, assuming a lumped capacitance response. A smaller time constant allows the system to return to equilibrium more rapidly and results in a faster detector response.

2.4.1 Thermistor Bolometers

In a bolometer a temperature increase ΔT due to absorbed thermal energy produces a change in the electrical resistance R_o of the detector according to

$$\Delta R = R - R_o = R_o\,P\,\Delta T\ , \tag{2.12}$$

where P (K^{-1}) is the temperature coefficient of resistance. If the bolometer is connected in series with a voltage source, V_b, and a resistance R_L much larger than R_o, as shown in Figure 2.5, the voltage V across R_L varies with the temperature according to

$$\Delta V = \frac{V_b \, R_o \, P \, \Delta T}{R_L} \tag{2.13}$$

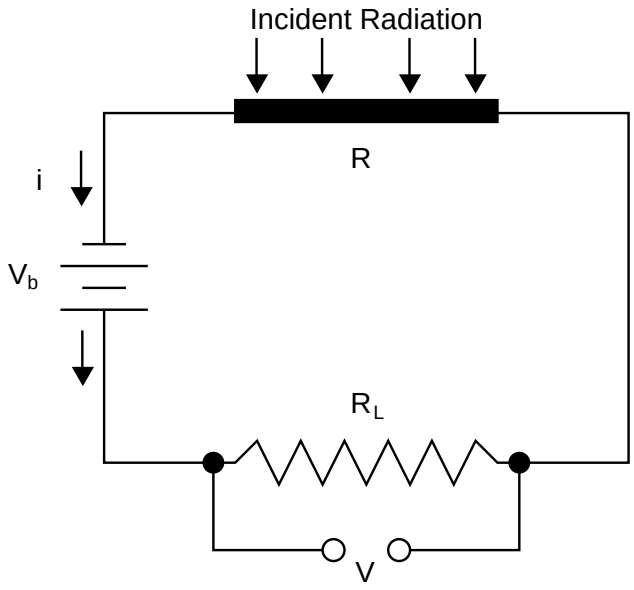

Figure 2.5: Schematic diagram of a typical bolometer circuit (adapted from Lenoble [1993])

A thermistor is a highly temperature-dependent resistor constructed of a semiconducting material. The resistivity of a thermistor is given by

$$\rho = \rho_o \exp\left[B\left(\frac{1}{T} - \frac{1}{T_o}\right)\right] \tag{2.14}$$

where ρ_o is the resistivity at a reference temperature T_o and B is a material characteristic of the thermistor. It should be noted that this relationship represents a negative coefficient of resistance, i.e., that the resistivity of the thermistor decreases with increasing temperature.

When used as an infrared detector, the thermistor is coated with a thin layer of black paint to maximize absorption of incident radiation, and is mounted on a heat sink to dissipate both

the absorbed radiant energy and the heat generated due to the current from the bias voltage across the thermistor. A thin layer of an insulating material is typically placed between the thermistor and the heat sink to prevent an electrical short-circuit and to increase the temperature rise and thus the reponsivity of the detector. The price paid for the increase in responsivity, however, is a decrease in the time response.

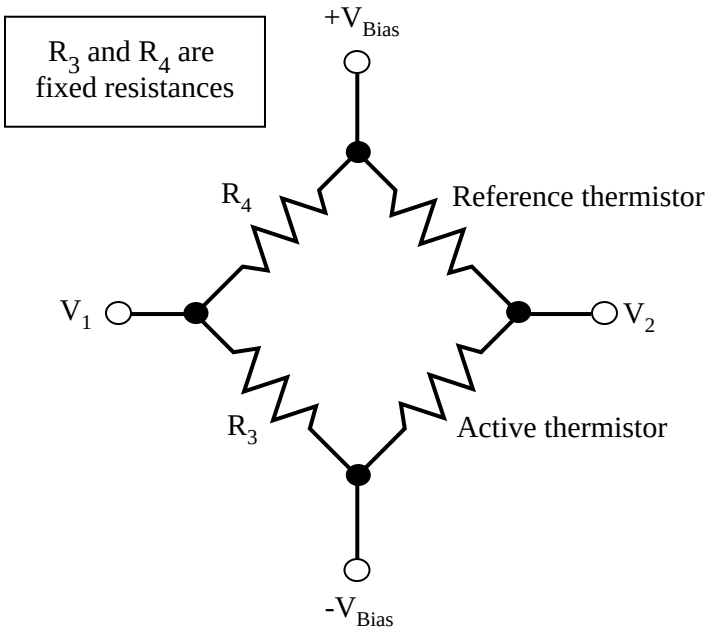

Figure 2.6: Wheatstone bridge configuration for a thermistor bolometer

The thermistor detector is usually placed in adjacent arms of a Wheatstone bridge circuit with a similar thermistor detector used as a reference detector, as illustrated in Figure 2.6. The reference detector tracks the instrument temperature while the active detector is allowed to vary with varying exposure to incident radiation. As the temperature of the active thermistor rises due to the absorbed energy, its overall resistance decreases, and a bridge voltage $V_{bridge} = V_2 - V_1$ is produced. A reference thermistor is used in the adjacent arm instead of a fixed resistor to compensate for any thermal drift in the temperature of the heat sink to which both are attached. For a more thorough treatment of thermistors, the reader is directed to Astheimer [1984].

2.4.2 Thermopiles

Thermopiles operate on the thermoelectric principles on which thermocouples are based. The three thermoelectric effects that influence the behavior of the thermocouple are the Peltier, Seebeck and Thomson effects [Pollock, 1985].

The Seebeck effect, named for German scientist Thomas Johann Seebeck, states that an electric current will flow in a closed circuit composed of two dissimilar metals if their junctions are maintained at different temperatures. If the circuit is open, an electromotive force, E_{AB} is created. This emf, termed the relative Seebeck voltage, is directly proportional to the temperature difference ΔT between the two junctions,

$$E_{AB} = S_{AB}\,\Delta T \tag{2.15}$$

where S_{AB} (V/K) is the Seebeck coefficient. The Seebeck effect is the contolling factor in the performance of a thermopile, and Pollock [1985] maintains that the Seebeck effect is the result of both the Peltier and Thomson effects.

When used as thermal radiation detectors, one junction of the thermocouple is maintained at a reference temperature equal to the ambient temperature, T_a, while the other junction is typically coated with a black absorber material and allowed to store thermal energy by absorbing incident radiation. A thermopile consists of several thermocouples connected in series to increase the emf output. For n junction pairs, the output signal is

$$Y = \Delta V = n\,S_{AB}\,\Delta T\ (V) \tag{2.16}$$

Huang [1990] presents a formal development of Equation 2.16 as well as excellent methods for calibration and signal conversion of thermopiles for temperature measurements.

2.4.3 Pyroelectric detectors

A pyroelectric detector consists of a crystal that, when heated by absorbed radiant flux, produces electric charges on opposite faces due to a change in average position of the electrons in the crystal. This leads to a current in an external circuit only if the input signal is modulated. This is because the charge separation related to a temperature change quickly drains off through the finite electrical resistance of the crystal material.

2.5 Radiometric calibration

The following material has been adapted from a CERES Science Team report [CERES Science Team, 1995] and is presented to familiarize the reader with some important terms and concepts in the calibration of radiometric instruments.

A radiometric instrument is typically calibrated by illuminating the aperture with radiance from a blackbody source maintained at various temperature levels. At each level associated with a particular blackbody temperature, the sensor output is recorded and the total broadband filtered radiance ($\mathrm{W/m^2 \cdot sr}$) absorbed by the detector, I_f, is calculated from knowledge of the instrument spectral response function, $S(\lambda)$. A full-field uniform blackbody source that completely fills the instrument field-of-view is used. The total absorbed filtered radiance can be expressed as

$$I_f = \int_{\lambda_1}^{\lambda_2} I_s(\lambda)\, S(\lambda)\, d\lambda \tag{2.17}$$

The unfiltered radiance from the blackbody source, I_s, can be calculated from the spectral emissivity $\epsilon(\lambda)$ of the blackbody and Planck's blackbody radiation distribution function,

$$I_s(\lambda) = \epsilon(\lambda) I_{bb}(\lambda) = \epsilon(\lambda) \frac{C_1}{\lambda^5 \exp(\frac{C_2}{\lambda T} - 1)} \tag{2.18}$$

where C_1 and C_2 are the well-known first and second radiation constants.

The spectral response function of the instrument is defined as the fraction of energy incident to the instrument aperture that is absorbed by the instrument detector at each discrete wavelength. It is a function of all optical components of the instrument, such as filters, mirrors and blackened coatings, which transmit, reflect, or absorb the spectral energy. The theoretical spectral response $S(\lambda)$ for a two-mirror system can be expressed as

$$S(\lambda) = \tau_f(\lambda)\,\rho^2(\lambda)\,\alpha_b(\lambda) \tag{2.19}$$

where $\tau_f(\lambda)$ is the combined transmittance of the bandpass filters, $\rho(\lambda)$ is the reflectivity of the two silvered mirrors, and $\alpha_b(\lambda)$ is the absorptivity of the "black" coating on the detector. The fact that this approach does not take into account diffuse emissions or reflections from other objects that may lie in the optical path of the instrument, combined with the difficulty in accurately characterizing each individual component, has led to the spectral response function being typically determined experimentally. The characterization is typically accomplished by illumination of the instrument with monochromatic radiation at specific wavelengths and normalization of the response with that of a spectrally flat reference detector.

For a linear instrument, the broadband filtered radiance, I_f , is directly proportional to the instrument output. This relationship is expressed by

$$I_f = G(m - m_s) + I_s \tag{2.20}$$

where m is the measured instrument output when viewing the blackbody, m_s is the measured output when viewing the reference source (typically a blackbody source at a temperature much lower than the temperature range used for the calibration points), I_s, is the filtered radiance from the reference source, and G is the gain of the instrument. Alternatively, the gain can be expressed

$$G = \frac{I_f - I_s}{m - m_s} \tag{2.21}$$

The gain is calculated by using a least-squares analysis to regress the difference in absorbed filtered radiances from the blackbody and reference sources with the difference in measured instrument output when illuminated by the two sources.

The calibration procedure is a statistical regression of absorbed filtered radiances and measured instrument output, and the calculated gain can be used to predict the filtered radiance from a new measured instrument output. However, some uncertainty enters into this calculated gain due to noise in the measurements as well as uncertainties in other components of the calibration system. This translates into an uncertainty in the predicted filtered radiance obtained from the gain. One measure of this uncertainty is the *fidelity interval* [Barkstrom, 1998], which is the difference between the upper and lower bounds on the estimated radiance, I_U and I_L, obtained from the gain. The fidelity interval is related to the inverse confidence interval of the regression, in that while the confidence interval gives the upper and lower bounds of the estimated radiance in terms of the measured instrument output, m, the fidelity interval gives them in terms of the estimated radiance, $\hat{I}$, itself.

For a well-determined linear system with N calibration temperature levels and R calibration points at each level, the fidelity interval can be expressed as [Barkstrom, 1998]

$$\begin{pmatrix} I_U \\ I_L \end{pmatrix} = \hat{I} \pm t(\nu, 1-\alpha/2)\, G\, s \sqrt{1 + \frac{1}{NR} + \frac{(\hat{I} - \bar{I})^2}{S_{LL}}} \tag{2.22}$$

where $\hat{I}$ is the estimated filtered radiance for an observed instrument output, t is the Student-t statistic for ν degrees of freedom at the $1-\alpha/2$ significance level, G is the instrument gain as expressed in Equation 2.21, S_{LL} is the spread over all calibration levels according to

$$S_{LL} = \sum_{n=1}^{N} (I_n - \bar{I})^2 \; . \tag{2.23}$$

and s is the standard deviation calculated from

$$s = \sqrt{\frac{\sum_{r=1}^{R} \sum_{n=1}^{N} (\Delta m_{nr} - \Delta \bar{m}_n)^2}{NR - 2}} \tag{2.24}$$

In this chapter a brief overview of radiometric instruments has been presented. A description of two-mirror optical systems typical of space-based remote sensing instruments has been given, with a discussion of the parameters that characterize the performance of the optical system. Likewise, a brief description of the three most common thermal detectors used in these applications and their associated parameters has also been presented. Finally, a brief overview of radiometric calibration theory was provided, along with relations to estimate the uncertainty associated with radiometric measurements. These concepts have been provided to familiarize the reader with terms and parameters used in discussion of the radiometric modeling and design environment and associated results.

Chapter 3

Genetic Algorithms

Genetic algorithms, a special calss of evolutionary algorithms, provide an attractive tool to integrate with the numerical instrument models to develop an instrument design tool. Evolutionary algorithms are a special class of search algorithms based on biological processes in nature, such as selection, crossover, and mutation. Evolutionary algorithms are based on a population of individuals, with each individual representing a set of the problem parameters. Each individual is assigned a score based on how well it performs in terms of an objective function. These individuals compete and evolve based on probablilistic rules that vary between different available algorithms, but the basic idea is that the better solutions have a greater chance of surviving and generating offspring, with the overall quality of the population increasing with each successive generation, and the entire population evolving towards an optimal solution.

Several advantages may be associated with using evolutionary algorithms in search/optimization problems:

1. They are zeroeth-order search algorithms, meaning they do not rely on numerical estimations of derivatives of any functions. This is important in solving problems with nonlinearities and/or discontinuities [Goldberg, 1989]

2. They are extremely robust algorithms that can be easily applied to a variety of different problems with good results [Schwefel, 2000]

3. Since they are population-based methods, they can search many different areas of the searchspace simultaneously, which decreases the chance of premature convergence on a local minima [Goldberg, 1989]

4. They have the property that they can obtain an optimal solution working with approximate function evaluations [Grefenstette and Fitzpatrick, 1985; Beasley et al., 1993a].

However, evolutionary algorithms have disadvatages as well. Because they are population based, they require many function evaluations. This can be a problem if the objective function is computationally expensive. Also, they tend to be slow-finishing, in that they can get in the neighborhood of the optimal solution quickly, but can take many generations to evolve towards the solution with a high degree of accuracy [De Jong, 1992]. Finally, while evolutionary algorithms generally perform well on a variety of hard-to-solve problems, they can often be outperformed by an algorithm developed soley for a single problem [Beasley et al., 1993a]. Evolutionary algorithms are at their strongest solving search problems with noisy, multi-modal searchspaces, for performing multi-objective optimization, and for problems for which classic search techniques fail and/or problem-specific techniques have not been developed [Schwefel, 2000].

3.1 Genetic Algorithms

Genetic algorithms are a class of evolutionary algorithms developed by John Holland at the University of Michigan [Holland, 1975]. In genetic algorithms, each parameter of the problem is encoded onto a gene (or allele), and then all the genes are combined into a chromosome, such that for a problem containing N variables, each chromosome is made up of N genes. Holland's original work used binary encoding for the genes, but current genetic algorithms

use many different types of encoding, with binary and real number representations being the most prevalent [Gen and Cheng, 1997].

Genetic algorithms use selection, crossover, and mutation operators. A typical genetic algorithm will begin by initializing the population, usually by assigning random values to the genes in each chromosome. Each chromosome is applied to the objective function and assigned a fitness based on how well it performs. In an optimization setting, the objective function is typically the function or problem that is being optimized. The fitness value is typically the value of the objective function in a maximization problem, and for a minimization problem it is typically the inverse of the returned value or the value subtracted from a constant. Once a fitness has been assigned to each chromosome, the selection operator is applied to select pairs of chromosomes for reproduction. Reproduction itself is handled by two operators, crossover and mutation. Typically, crossover is applied to a set of parents to produce offspring, with each offspring having some probability of undergoing mutation. At this point the offspring are inserted into the population, with the rule governing insertion varying among the different types of GA's. Each indivdual is then evaluated, a fitness is assigned, and selection and reproduction occurs once more. This process, which is repeated until the algorithm reaches the stopping point, is illustrated in Figure 3.1.

One important aspect of GA's is that they are blind; that is, the algorithm has no way of knowing if it has arrived at the correct solution and must be told when to stop [Goldberg, 1989]. While at first sight this might seem a disadvantage, it is an important characteristic in preventing GA's from getting trapped on local minima. The most common stopping criterium are stopping after a preset number of generations or when the average fitness of the population has converged to the same value; i.e., when the entire population is clustered around some point in the searchspace.

This is a general overview of a generic genetic algorithm. The following sections contain more detailed descriptions of some of the operators and other issues involved with genetic algorithms.

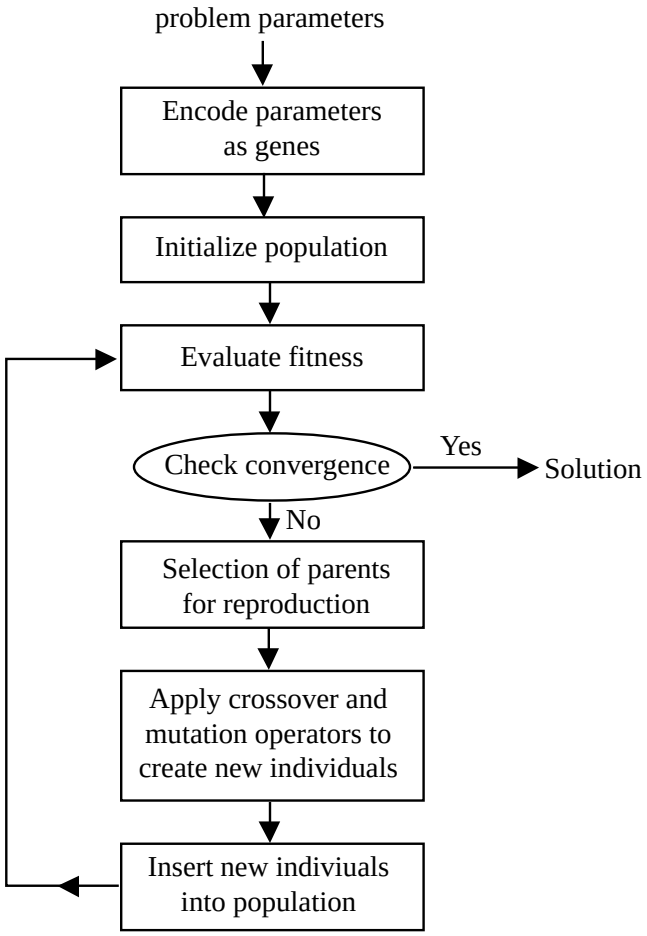

Figure 3.1: Flowchart for simple genetic algorithm

3.1.1 Encoding

As mentioned earlier, several different methods for encoding the search parameters exist, with binary and real numbers being the two most popular, at least for engineering and optimization applications. However, the question of which encoding to use is a major one, and no definitive answer exists [Antonisse, 1989]. The original genetic algorithms developed by Holland used binary encoding, and for years this was the only type of encoding used. In binary encoding, each gene is represented by a binary string of fixed length. The individual strings representing the genes are then concactenated into one long binary string that

represents the chromosome.

Other types of encoding have emerged as the use of genetic algorithms has spread to other disciplines outside of computer science; however, a school of thought exists that maintains that binary encoding is the only proper way to use genetic algorithms, and that they lose some of their usefulness with other types of encoding. Still, it has been shown that for real-valued continuous functions, using real-valued encoding to represent the genes significantly outperforms binary encoding [Davis, 1991; Wright, 1991; Eshelman and Schaffer, 1992]. In real-value encoding, each gene is represented by a real number, and the chromosome is represented by an array of genes.

The choice of the optimal encoding to use is extremely problem dependent [Davis, 1991]. One obvious limitation in using binary encoding is that it can only search discrete points in the searchspace, whereas a real-valued GA can find any point in the searchspace, within the limits of the computer's precision [Eshelman and Schaffer, 1992; Wright, 1991]. Using binary strings, the searchspace can be resolved into a mesh of any desired precision, but higher precision can result in extremely long strings. Another issue is the decoding of the chromosome by the objective function. The objective function must be able to use the information contained in the chromosome to return a fitness value. If the objective function is real-valued and continuous, it is more efficient to use an array of real-numbers to represent the problem parameters [Wright, 1991]. However, if the objective function is by nature discrete, such as a problem containing many different possible configurations, then a binary encoding would be more appropriate. Another factor to consider is that, since genetic algorithms were initially developed using binary encoding, far more research and literature exists on GA's using binary encoding than any other type of encoding. In fact, almost all of the mathematical theory concerning genetic algorithms, including Holland's Schemata Theory [Holland, 1975] and Goldbererg's Building Block Hypothesis [Goldberg, 1989], are based on a binary encoding of the problem. Little published information exists on specific operators and techniques for other types of encoding.

3.1.2 Fitness

While several parameters can influence the performance of a genetic algorithm (population size, type of selection, etc.), the most important parameters by far are the type of encoding used and the definition of the objective, or fitness, function [Beasley et al., 1993a]. Genetic algorithms are sufficiently robust to perform well under a variety of different settings, but encoding of a gene and the definition of a fitness function that accurately represents the problem of interest are essential [Davis, 1991].

The fitness of a chromosome is a numerical representation of how that chromosome performs in solution of the problem of interest. The fitness function should decode a chromosome, apply it to the problem, and return a value based on its performance. For example, if the problem at hand is to minimize the cost of a product under development, the fitness function could read a chromosome containing values for the design variables, such as dimensions, materials, etc., compute the cost, and return the cost as the fitness. It is fairly easy to define a fitness function for problems where one parameter (cost, error, etc.) is optimized againt several variables, as the parameter to be optimized can easily represent the fitness. More difficult are multi-objective optimization problems, where several parameters (cost, performance, time) must be optimized for many design variables. The fitness function must be constructed so that it returns one value that represents how well the chromosome optimizes all of these parameters. Several different techniques exist for this, and this is an area of ongoing investigation. The interested reader is directed to Schaffer [1985], Fonseca and Fleming [2000], and Hajela and Lin [1992].

Another consideration that impacts the fitness function is that of constraints. Since the genetic algorithm contains abstract encodings of the problem parameters, it is difficult to apply problem constraints directly in the algorithm. Instead, if a chromosome contains one or more parameters that violate a problem constraint, that chromosome is accorded a very low fitness value automatically, according to some type of penalty function [Richardson et al., 1989]. For example, in a parameter estimation problem where negative parameters are

impossible, a chromosome that contains a negative value could be assigned a fitness value of zero. Chromosomes with a very low fitness value will have a poor chance of reproducing and will eventually be driven from the population. An advantage to this method is that the level of fitness assigned can be varied by the strength of the constraints, thus allowing for some exploring outside the constraints with a penalty, if desired [Michalewicz and Janikow, 1991].

3.1.3 Selection

One of the primary operators that allows a genetic algorithm to evolve towards a good solution is *selection* . Selection is the process of deciding which individuals in the population will be allowed to mate to produce offspring. This is a very important step in determining the convergence of the GA [Miller and Goldberg, 1995]. If too many individuals with low fitness are allowed to mate, the population will evolve slowly, if at all. However, if too many extremely fit individuals are allowed to mate, they can in effect take over the population and cause premature convergence, i.e., convergence on a local solution. A balance must be struck between fit individuals driving the evolution of the population with less fit individuals remaining to maintain genetic diversity in the group.

Selection is usually accomplished by allowing each individual some chance of being put into an intermediate mating pool, with fitter individuals having a greater chance of having multiple copies in the mating pool, and less fit individuals having decreased chances of having even a single copy. Parents are then chosen at random from this mating pool. Individuals can be chosen for the mating pool in many different ways. The most common include fitness scaling, fitness ranking, and tournament selection.

Fitness scaling puts a certain number of copies of individuals into the mating pool, with the number being proportional to the fitness. This results in the probability of an individual being selected to reproduce being directly proportional to its fitness. The problem is that a superfit individual with an extremely high fitness can put so many copies in the mating

pool that all others have an extremely small chance of being selected to mate [Hancock, 2000]. A better approach is *fitness ranking*. Here the population is sorted and ranked based on fitness. Each individual is given a predetermined number of copies based on its rank. The advantage here is that the highest ranked individual has a set number of copies in the mating pool, no matter how high its fitness is relative to the other individuals. This prevents a superfit individual from dominating the population [Grefenstette, 2000]. Another commonly used method is *tournament selection*, where individuals are chosen at random and the best n individuals are selected to go into the mating pool [Blickle, 2000]. The simplest variant is the *binary tournament*, where two individuals are chosen at random and the fitter is assigned to the mating pool. A tournament selection pressure can be applied by letting the less fit individual be copied instead a certain percentage of the time. This helps to maintain diversity in the mating pool [Goldberg and Deb, 1991].

3.1.4 Crossover

Along with selection, the operator that probably most impacts the performance of a genetic algorithm is *crossover*. A genetic algorithm performs by two primary methods: exploration and exploitation [Spears, 1992]. Exploitation refers to exploiting existing points in the searchspace to investigate those areas in the searchspace that lie within a hypersphere made up of all the existing points. Exploration refers to the investigation of new points that lie outside of this hypersphere. Exploitation is achieved by crossover, although some points within the hypersphere can be reached by mutation as well, while exploration can only be achieved by mutation [Whitley, 1993].

Crossover is the creation of one or more offspring, or children, from the genetic material of two parents chosen with the selection operator. Crossover is achieved by blending the genes from each parent to create a child that has some of the characteristics of each parent. In an evolutionary algorithm, many different methods can be used to achieve this, with these methods varying by the type of problem and most significantly, with the type of encoding.

The crossover methods for binary encoding bear a more direct connection to that seen in nature. Recall that in binary encoding, the chromosome consists of one long binary string, made up of substrings containing the individual genes. The simplest crossover technique, introduced by Holland [1975], is known as *one-point crossover* . In this technique, two parent chromosomes of equal length are selected, a random point on the chromosome is selected as the cut-point, and the offspring chromosome is formed from the binary string up to the cut-point from one parent, and past the cut-point from the other parent. For example, if we have two chromosomes, the first being represented as 11111 and the second as 00000, and we randomly select a cut-point after the third position on the chromosome, then the resulting child would be either 11100 or 00011. A common technique is to form two children from each crossover using both possibilities, but this is not necessary.

Two-point crossover is similar to one-point crossover, except that, as the name implies, two cut-points are used. The offspring contains the string between the two cut-points from one parent, and the strings outside the two points from the other parent. Using the two parent chromosomes from the above example, if the two cut-points were located after the second and fourth positions on the chromosome, the resulting offspring would be either 00110 or 11001. Thus, one-point crossover is simply a specialized form of two-point crossover, with one of the cut-points being the beginning or end of the chromosome. It is generally agreed upon by researchers that two-point crossover provides superior results to one-point crossover [Schaffer et al., 1989].

Another popular crossover technique used for binary encoding is *uniform crossover* . In this technique, each bit on the offspring is taken from one of the two parents according to what is known as a crossover mask. The crossover mask is a randomly generated binary string of the same length as the chromosomes. For each space on the offspring, if the corresponding space on the crossover mask is a 1, the bit from the first parent is applied to the offspring, and if it is a 0, then the bit is taken from the second parent. A new crossover mask is generated randomly for each new offspring [Syswerda, 1989].

These three techniques are widely used in practice, but several other techniques exist. No general consensus exists as to which is the best technique, with two-point, uniform, and mutiple-point crossover all having their proponents. This is an area of vigorous ongoing debate and research [Beasley et al., 1993b].

Unlike the selection operator, which operates in the same manner regardless of binary or real encoding, the cross-over operator for real encoding is quite different. The two parent chromosomes, each consisting of an array of real numbers, form the offspring by performing some type of arithmetic operation on each number in the array [Booker et al., 2000]. Usually this is some type of averaging. The simplest technique is to form each gene of the offspring, represented by one number in the array, by averaging the corresponding genes from the two parents [Wright, 1991]. This is repeated for each gene until the new chromosome is formed. Another common approach is to perform a weighted average of the genes using the fitness of each parent. The geometric mean, which is the square root of the product of the two genes, is also sometimes used [Michalewicz et al., 1996].

Aside from averaging, other approaches exist for selecting a point lying in the searchspace between the two parents. The first is completely arithmetic crossover. This is done using a simple equation

$$C_1 = RP_1 + (1 - R)P_2 \tag{3.1}$$

where P_1 and P_2 are the two parent chromosomes and R is a uniformly distributed random number between 0 and 1 [Michalewicz, 1996]. Another method is to average the two parents but add or subtract a random component. This method was proposed by Doyle [1995] and is summarized as

$$C_1, C_2 = \frac{P_1 + P_2}{2} \pm |P_1 - P_2|R \tag{3.2}$$

where again P_1 and P_2 are the two parent chromosomes and R is a uniformly distributed random number between 0 and 1. This can be used to generate two offspring at a time, or can be modified to generate one offspring from two parents by randomly selecting between adding or subtracting the random component. Another method is to either add the difference between the two genes to the higher value or subtract it from the lower value. This method is known as extension and is unusual in that it explores the area in the searchspace outside the hypersphere formed by the two parents [Eshelman and Schaffer, 1992].

3.1.5 Mutation

As stated earlier, *mutation* is the method by which the genetic algorithm explores new areas of the searchspace. Studies have shown the performance of genetic algorithms to be more sensitive to mutation rates than crossover rates [Schaffer et al., 1989]. Mutation is an operator that acts on a single chromosome, changing the value of a gene on that chromosome slightly. How this is achieved depends on the type of encoding used. In the case of binary encoding, a random bit on the string is "flipped", meaning a 0 is changed to 1, and vice versa [Goldberg, 1989].

For the case of real encoding, the method can be more complex. The idea is to perturb one of the genes to create a new chromosome [Wright, 1991]. One commom method is to generate a random number with a Gaussian distribution having as a mean the unperturbed value of the gene and a variance specified by the user [Back et al., 2000]. This technique has the advantage of generating most new points in an area close to the original point, but having a small probability of generating points far from the original point, which helps to explore new areas.

One point that should be discussed here is the slightly different role mutation can play when real encoding is used. In binary encoding, there exists a discrete number of points, and mutation is essentially used to jump to a random point in the searchspace, which may or may not be near the original point. Since there exist a finite number of points, this type of

random jump in conjunction with crossover explores the searchspace effectively. However, in real encoding, where the searchspace is continuous, the mutation operator is sometimes used both to perform a kind of local search around the original point, and in its more traditional role of generating a random point in the searchspace. This requires two mutation operators, which are sometimes known as *creep mutations* and *random (jump) mutations* [Beasley et al., 1993b]. As the name implies, creep mutation refers to a mutation of a real-valued chromosome where the numerical value of one or more genes is changed by adding or subtracting a small randomly generated number. *Geometric creep* is a variation where the value of the gene is multiplied by a random number close to unity. Jump mutation is an operator that replaces the numeric value of one or more genes on the chromosome with a value selected randomly from the searchspace.

It should be noted that, since operators for real-valued chromosomes consist of arithmetic operations applied to the numerical values contained within the chromosome, a plethora of variations on the basic operators given here exist. Particularly for crossover, the operators can be as simple or complex as desired. The basic operators presented here, however, will handle most problems reasonably well.

3.1.6 Replacement Methods

Once the selection, crossover, and mutation operators have been applied and a suitable number of new chromosomes have been created, the final challenge is finding the best method to insert them into the population. These methods fall into one of two main categories: *generational*, or *simple genetic algorithms*, and *steady-state genetic algorithms* [Sarma and De Jong, 2000].

In a simple GA, the entire population is replaced by the offspring every generation. In order to maintain the population size, a number of offspring equivalent to the current population size must be created each generation. While simple GA's offer a rapid increase of population fitness each generation, replacing the entire population each generation runs the risk

of throwing out good solutions before they have been fully exploited. To counter this, a technique known as *elitism* is often employed. In elitist simple GA's, the N best individuals are copied directly into the next generation and the balance of the next generation is filled with individuals created by crossover and mutation. This strategy provides the exploitation and exploration of crossover and mutation while retaining the best current solutions [De Jong, 1992].

A steady-state GA, on the other hand, only generates a certain number of new individuals, and replaces an equal number of individuals in the current population to form the next generation. Deciding which individuals to replace in the current population is a key factor in determining the performance of a steady-state GA. Some methods include replacing the N worst individuals or replacing the individual most similar to the new individual, to maintain diversity in the population [Whitley, 1989; Syswerda, 1989]. While steady-state GA's will evolve more slowly (on a generational basis) than simple GA's, they can more fully exploit current solutions [Beasley et al., 1993a]. Also, a factor that should not be discounted is computation time; if a steady-state GA only generates a number of new individuals equal to 10 percent of the population size, then the number of function evaluations per generation would be one-tenth the number of an equivalent sized simple GA.

3.1.7 Finishing Criteria

One advantage to the use of the genetic algorithms is that they are, in essence, blind to the problem. This is because they work with an abstraction of the problem contained in the encoding and fitness values. This is useful because it makes genetic algorithms extremely robust and general over a wide range of problems. One disadvantage, perhaps, is that they do not know when to stop without being explicitly told. The GA is typically allowed to run for a predetermined number of generations sufficient for convergence, or until the average fitness of the population converges on some value, indicating that all members of the population are clustered around the same point in the searchspace.

A convergence criterion based on the problem to be solved can be used, but this may not be desirable. One problem is that this may require *a priori* knowledge of the problem and its solution. The second is that removing some of the abstraction between the genetic algorithm and the problem at hand reduces the algorithms robustness and generality, by transforming it into a problem-specific algorithm [Goldberg, 1989].

3.1.8 Hybrid Genetic Algorithms

The performance of genetic algorithms can be improved for some problems by incorporating existing domain knowledge or specific algorithms to create hybrid genetic algorithms [Michalewicz, 2000]. This takes advantage of the robustness of the genetic algorithm while utilizing existing methods specific to the problem [Davis, 1991].

One approach is to incorporate specific knowledge of the problem into the objective function or constraints via penalty functions. Another is to incorporate neighborhood search algorithms, such as direct-descent or simplex algorithms. This can be done both during or after execution of the genetic algorithm routine. Local search operators can be used to search the neighborhood surrounding each new individual produced from cross-over, with the optimal local point replacing the individual prior to its evaluation in the objective function. This is one method of inserting Lamarkian learning into the genetic algorithm [Gen and Cheng, 1997]. Another approach is to run the genetic algorithm with its normal stopping criteria (number of generations or average fitness) and then use the result as a starting point for a problem-specific or local search method. While this method can be used to overcome the slow-finishing aspects of GA's, it could increase the possibility of premature convergence on a suboptimal point.

This chapter serves as an overview of genetic algorithms, and introduces the reader to the basic algorithm and some of the specific mechanisms used, such as selection, crossover, mutation, and replacement. In addition, a discussion of two types of encoding, floating-point and binary, is presented along with and the relative merits of each. The objective

function is also discussed, along with its role in representing the problem of interest, such as an engineering optimization problem, in the GA environment. In the following chapters, an end-to-end optical and electrothermal radiometric instrument modeling environment will be presented in detail, as well as a description of its use with GA's to optimize the model's performance and to aid in the development of new instrument concepts.

Chapter 4

Radiometric instrument modeling environment

A radiometric instrument modeling environment has been developed in an attempt to aid in the analysis and design of radiometric instruments. The motivation for developing this tool is the need to simplify the modeling of radiometric instruments by eliminating the need to create a dedicated Monte Carlo ray-trace model and/or a dedicated electrothermal model for each standard radiometric instrument configuration. Also, modeling avoids the need to build and test expensive prototypes. The environment has been developed as a series of C++ classes and programs that allow the user to develop coupled optical and electrothermal models of the instrument optical and detector subsystems. A schematic digram of the model is presented in Figure 4.1. The user can create an integrated end-to-end model of a radiometric system by specifying various parameters that define the optical and detector subsystems. A Monte Carlo ray-trace (MCRT) model of the optical system and a coupled finite-difference electrothermal model of the detector system results. The model provides a dynamic simulation of the instrument's behavior, using a dynamic radiative input to the instrument to produce the time varying electric signal. In addition, various other parameters that characterize the instrument's behavior can be produced, such as the distribution

of absorbed energy and transient temperature distributions in the detector assembly, among others.

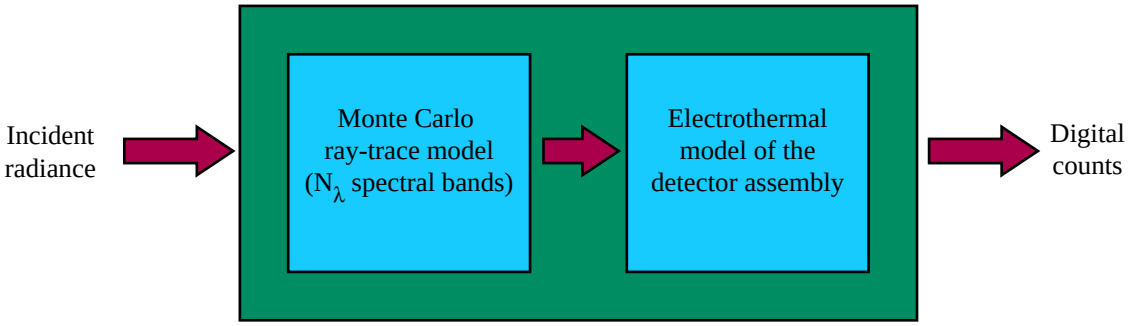

Figure 4.1: End-to-end integrated radiometric instrument model

4.1 Optical model

The optical model described in this chapter uses the Monte Carlo ray-trace (MCRT) method to determine the distribution of thermal radiation throughout the instrument housing and particularly, on the surface of the detector. The current version of the optical model is restricted to a two-mirror configuration, as this is the most common configuration in space-based remote sensing applications, which were the driving force behind model development. The optical model is fully spectral, and reads the spectral absorptivity and specularity ratio values and the associated wavelength intervals for each participating surface from an input file. The assumption of gray surfaces can be used by defining only one wavelength interval over the spectral range of interest and providing the appropriate values for absorptivity and specularity.

The optical model allows the user to select from five standard two-mirror configurations:

1. Cassegrain (parabolic primary mirror, hyperbolic secondary mirror)
2. Gregorian (parabolic primary mirror, ellipsoid secondary mirror)

3. Dall-Kirkham (ellipsoid primary mirror, spherical secondary mirror)

4. Ritchey-Chrétian (two hyperbolic mirrors)

5. Modified Cassegrain (two spherical mirrors)

Each mirror can be defined by the location of its vertex on the optical axis, Z_v, the outer radius of the mirror slice, a, and the mirror depth, c. These dimensions are illustrated in Figure 4.2. Note that this is consistent with the notation used by Walkup [1996].

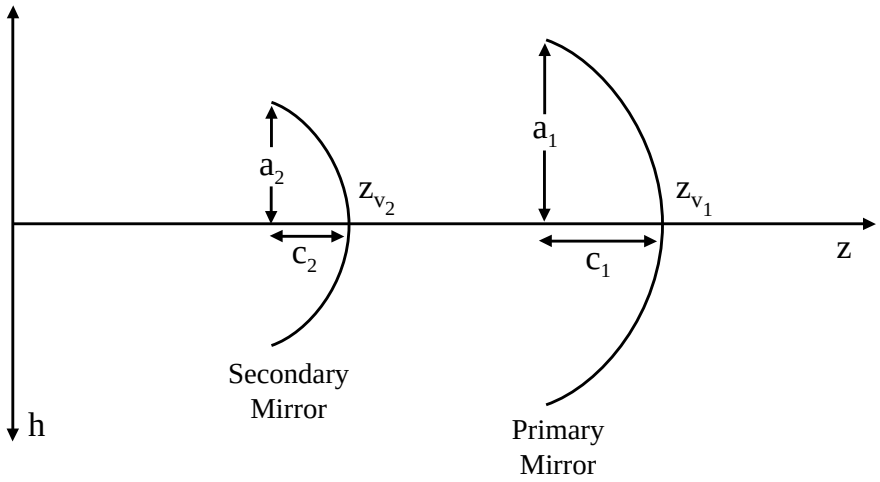

Figure 4.2: Illustration of the optical coordinate system and mirror input parameters

The following sections provide a description of the four conic surfaces used as reflectors in the modeling environment: sphere, parabaloid, hyperboloid, and ellipsoid. The analytic equations for each surface and their relations to the input variables a, c, and Z_v are provided. The analytic expressions were adapted from Walkup [1996] and are in agreement with those provided by Korsch [1991].

4.1.1 Optical theory

This section provides a brief review of geometric optics used in the radiometric instrument modeling environment, including mathematical descriptions of the various conic surfaces used as mirrors, an overview of Gaussian (first-order) optical theory, and a description of the quadric surfaces used in the model.

For the radiometric instrument modeling environment, we consider only mirrors that are symmetrical about the optical axis, which coincides with the z-axis in the global coordinate system. We can then calculate all mirror parameters using the meridonial coordinate system, in which

$$x^2 + y^2 = h^2 \tag{4.1}$$

as illustrated in Figure 4.2.

The curvature of a mirror, ρ, is related to the vertex radius of a mirror, R, by

$$\rho = \frac{1}{R} \tag{4.2}$$

where R is a characteristic of each conic shape. The *deformation constant* is a parameter used to characterize the performance of an optical system, and is calculated from the two focal points of an axisymmetric function representing a surface by

$$\delta = -\frac{f_1 - f_2}{f_1 + f_2} \tag{4.3}$$

The object and image planes of a given optical surface are related by

$$\frac{n'}{l'} - \frac{n}{l} = \rho(n' - n) \tag{4.4}$$

where l, l' represent the distance from the optical surface to the object and image plane, respectively, and ρ represents the surface curvature [Mouroulis and MacDonald, 1997]. Distances to the left of the vertex are considered negative and those to the right are considered

positive. For the case of a mirror, $n = n'$, and thus Equation 4.4 reduces to

$$\frac{1}{l'} + \frac{1}{l} = 2\rho = \frac{2}{R} \tag{4.5}$$

In the following paragraphs we explore the ramifications of Equation 4.5 applied to spherical, parabolic, hyperbolic, and elliptical mirrors.

Spherical mirrors

The geometric equation for a sphere, in a local x, y, z-coordinate system, is

$$x^2 + y^2 + (z - z_o)^2 = h^2 + (z - z_o)^2 = R^2 \tag{4.6}$$

and the radius of the sphere is given by

$$R = \frac{a^2 + c^2}{2c} \tag{4.7}$$

where a and c are the mirror width and depth and correspond to the input dimensions to the optical modeling environment. The vertex radius of a sphere is equal to its true physical radius. The deformation constant is zero for a sphere ($\delta = 0$) and is generally referred to as the aspheric deformation constant.

Both focal points of a sphere lie at its geometric center, a distance R from the mirror vertex, that is

$$f_{1,2} = R \tag{4.8}$$

Parabolic mirrors

The geometric equation for a parabaloid is

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{(z - z_v)^2}{c} \tag{4.9}$$

where $a = b$ due to the symmetry about the optical axis. The vertex radius can be calculated from the input mirror dimensions a and c using

$$R_v = \frac{a^2}{2c} \tag{4.10}$$

The aspheric deformation constant is negative unity for a parabaloid ($\delta = -1$). The two focal points are given by

$$f_1 = \infty, \; f_2 = \frac{R_v}{2} \tag{4.11}$$

Hyperbolic mirrors

The hyperboloid is the only conic section that consists of two separate surfaces. Its geometric equation is given by

$$\frac{(z - z_o)^2}{{c_h}^2} - \frac{x^2}{A^2} - \frac{y^2}{B^2} = 1 \tag{4.12}$$

The coefficient c_h is the distance between the center of the two sheets and each vertex, and is given by

$$c_h = \frac{c}{\sqrt{2} - 1} \tag{4.13}$$

The coefficients A and B are set equal to each other due to symmetry about the optical axis. As noted by Coffey [1998], the coefficient A is not the same as the coefficient a used to describe the width of the mirror slice, as erroneously stated by Walkup [1996], but can be related to the input dimensions by

$$A = \frac{ac_h}{c_h + c} \tag{4.14}$$

The vertex radius can be calculated from

$$R_v = \frac{A^2}{c_h} \tag{4.15}$$

the aspherical deformation constant is given by

$$\delta = -\left(\frac{\sqrt{A^2 + {c_h}^2}}{c_h}\right)^2 \tag{4.16}$$

and the focal points are given by

$$f_{1,2} = \sqrt{A^2 + {c_h}^2} \pm c_h \tag{4.17}$$

with one lying on either side of the mirror vertex.

Elliptical mirrors

The geometric expression for an ellipsoid is

$$\frac{x^2}{A^2} + \frac{y^2}{B^2} + \frac{(z - z_o)^2}{C^2} = 1 \tag{4.18}$$

where

$$A = B < C \tag{4.19}$$

due to symmetry and to obtain a prolate ellipsoid. Unfortunately, the minor axis coefficient A is not directly related to the input dimensions a and c, and must be separately specified. The major axis coefficient can then be specified by

$$C = \frac{c}{1 - \sqrt{1 - \dfrac{a^2}{A^2}}} \tag{4.20}$$

the vertex radius is given by

$$R_v = \frac{A^2}{C} \tag{4.21}$$

the aspherical deformation constant is given by

$$\delta = -\left(\frac{\sqrt{A^2 - C^2}}{C}\right)^2 \tag{4.22}$$

and the two focal points are given by

$$f_{1,2} = \sqrt{C^2 - A^2} \pm C \tag{4.23}$$

4.1.2 Monte Carlo ray-trace method

The Monte Carlo ray-trace (MCRT) method is a statistical approach to modeling radiation heat transfer problems [Mahan, 2002]. The radiative exchange between the participating surfaces is treated as the summation of the effects of individual energy bundles being emitted, reflected, and absorbed by the various surfaces. In the MCRT method, the paths of these energy bundles are called "rays." Each energy bundle is emitted from a surface in a random direction, and its path traced until it intersects a participating surface. A probabilistic approach is used to determine if the energy bundle is absorbed or reflected, where the absorptivity of the surface can be interpreted as the probability that an individual energy bundle striking that surface will be absorbed. If the energy bundle is reflected, a probabilistic method is used to determine the direction of reflection, and the energy bundle is traced until it strikes another surface. When the energy bundle is finally absorbed, the location of absorption is recorded and the process is repeated for another energy bundle. This process is repeated for a large number of energy bundles being emitted from each surface. Typically, on the order of several million energy bundles are emitted from each surface to achieve adequate sampling.

MCRT methodology

The first step in a Monte Carlo ray trace is to determine the point of emission on the emitting surface for the individual energy bundle. For example, for a simple rectangular plane surface, the x and y locations can be obtained from

$$x = x_{min} + (x_{max} - x_{min})R_x \tag{4.24}$$

and

$$y = y_{min} + (y_{max} - y_{min})R_y \ , \tag{4.25}$$

where R_x and R_y are uniformly distributed random numbers drawn from the interval $R(0,1)$. The z coordinate can then be calculated from the equation of the surface, $S(x,y,z) = 0$. Similarly, for a plane disk, the emission point, in polar coordinates, can be calculated from [Walkup, 1996]

$$r = r_{max}\sqrt{R_r} \tag{4.26}$$

and

$$\phi = 2\pi R_\phi \ , \tag{4.27}$$

where r and ϕ are the radial and angular coordinates and R_r andR_ϕ are random numbers drawn from the interval $R(0,1)$.

The next step in the MCRT method is to determine the direction of emission. For the case of diffuse emission, the zenith and azimuth angles, θ_e and ϕ_e, can be determined from [Meekins, 1990; Sparrow and Cess, 1966]

$$\theta_e = \sin^{-1}(\sqrt{R_\theta}) \tag{4.28}$$

and

$$\phi_e = 2\pi R_\phi \ , \tag{4.29}$$

where R_θ and R_ϕ are sequential random numbers drawn from the interval $R(0,1)$. It should be noted that θ_e and ϕ_e are calculated in terms of the local coordinate system for the surface in question, and must be converted to values in the global coordinate sytem, as described in Mahan [2002].

Once the location and direction of emission are determined, the surfaces that the path of the emitted energy bundle will intersect must be determined. In the MCRT models described in this dissertation, all surfaces are modeled as quadrics. A quadric is a surface that can be represented by the parametric equation

$$Q = Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Gx + Hy + Iz + J = 0 \tag{4.30}$$

where A, B, C, D, and E are parameters whose values determine the nature of the surface.

A quadric can also be expressed in matrix form as

$$M_Q = \begin{bmatrix} A & D & E & C \\ D & B & F & H \\ E & F & C & I \\ C & H & I & J \end{bmatrix} \quad P = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{4.31}$$

For a point P to be contained on quadric surface Q, then

$$Q + P^T \cdot M_Q \cdot P = 0 \tag{4.32}$$

must be true. Each quadric surface making up an enclosure is checked to see if an intersection with the path of the energy bundle exists.

When the intersection between the energy bundle and a surface is found, then a probabilistic approach is used to determine if the enegy bundle is absorbed. A random number is drawn from a large population of uniformly distributed random numbers on the interval $(0, 1)$. If the random number is less than the absorptivity of the surface in question, the energy bundle is absorbed; otherwise the energy bundle is reflected. The reflectivity can be treated approximately as the sum of diffuse and specular components [Sparrow et al., 1962], according to

$$\rho = \rho^d + \rho^s \tag{4.33}$$

where ρ^d is the diffuse component of reflection, and ρ^s is the specular component of relection. The specularity ratio of the surface, r^s, is then defined

$$r^s = \frac{\rho^s}{\rho^d + \rho^s} \tag{4.34}$$

Although this model has practical limitations, it can be quite accurate if used in an enclosure where the reflectivity is small and the specularity ratio is large, as it is on interior surfaces of optical instruments such as cameras and telescopes. If the energy bundle is reflected, a random number is drawn, and if its value is less than that of the specularity ratio, the energy bundle is reflected specularly and the direction of reflection can be calculated from the direction of incidence; otherwise, the ray is reflected diffusely, and the direction of reflection is calculated as if the energy bundle were diffusely emitted with no knowledge of the incident direction. If the energy bundle is absorbed, a counter is incremented and a new energy bundle is emitted and traced through the enclosure until it too is absorbed. This process is repeated for N energy bundles, with N being sufficiently large to achieve satisfactory sampling.

4.1.3 Distribution Factors

The total diffuse-specular distribution factor D_{ij} between two gray surfaces is defined as the fraction of power emitted by surface i that is ultimately absorbed by surface j, both directly as as a result of all possible reflections [Mahan, 2002], according to

$$Q_{ij} = \epsilon_i A_i \sigma T_i^4 D'_{ij} \ (W) \tag{4.35}$$

where Q_{ij} is the energy emitted by surface i that is absorbed by surface j, ϵ_i, A_i, and T_i are the emissivity, area, and temperature, respectively, of surface i, and D'_{ij} is the diffuse-specular distribution factor from surface i to surface j.

The diffuse-specular distribution factor has three useful properties when dealing with gray

surfaces, summarized as

$$\sum_{j=1}^{n} D'_{ij} = 1.0 \ , \qquad i = 1, 2, ..n \tag{4.36}$$

$$\epsilon_i A_i D'_{ij} = \epsilon_j A_j D'_{ji} \ , \qquad i, j = 1, 2, ..n \tag{4.37}$$

and

$$\sum_{i=1}^{n} \epsilon_i A_i Dij' = \epsilon_j A_j \ , \qquad j = 1, 2, ..n \tag{4.38}$$

where Equation 4.36 is a statement of the conservation of energy, Equation 4.37 is a statement of reciprocity, and Equation 4.38 can be derived by combining the previous two equations.

The distribution factor can be estimated in the MCRT method by

$$D'_{ij} \cong \frac{N'_{ij}}{N_i} \tag{4.39}$$

where N'_{ij} is the number of rays emitted by surface i that are absorbed by surface j, and N_i is the total number of rays emitted by surface i.

For the special case of heat transfer between an opening in an enclosure and a surface within the enclosure, the distribution factor between the opening and surface j can be defined as the fraction of energy entering the enclosure that is absorbed by surface j according to

$$Q_{oj} = Q_o D_{oj} \tag{4.40}$$

where Q_o is the energy entering the enclosure through the opening, Q_{oj} is the fraction of that energy that is absorbed by surface j, and D_{oj} is the distribution factor between the opening and surface j. The distribution factor can be estimated by treating the opening as a virtual surface and firing rays from the surface and estimating D_{oj} according to Equation 4.39.

4.1.4 Monte Carlo ray-trace (MCRT) model

A model of the generic telescope optical system has been developed using the Monte Carlo ray-trace method. The Monte Carlo ray-trace method is a statistical method widely used in optics and thermal radiation analysis. In the MCRT method, a large number of energy bundles, or rays, is emitted from each emitting surface, and probabilistic methods are used to determine where they are ultimately absorbed.

For the purposes of the model formulated in this dissertation, the instrument aperture is considered to be the only emitting surface in the ray trace. Energy emitted from the instrument structure is treated as a constant heat flux on the detector surface. This is justified because the structure experiences minimal temperature variations during the duration of a scan, and the only variation in energy received by the detector is due to variations in energy entering the instrument aperture [Savransky, 1996]. Energy bundles are emitted from the aperture, either diffusely within a specified solid angle or at a specified azimuth and zenith angle, and tracked through the instrument until absorbed, and the spatial coordinates where they are absorbed on the detector surface are recorded. An estimate of the distribution factor between the aperture and the detector surface is then calculated using Equation 4.39 and used to convert the power entering the instrument aperture to absorbed power on the detector surface. The exact details of the ray-trace and distribution factors depend on whether the incident radiant intensity is spatially homogenous or non-homogenous, and are discussed in the relevant sections of this dissertation.

4.1.5 Spatially homogenous scenes

The assumption of a spatially homogenous scene is made when the magnitude of the incident radiance may be assumed equal from all angles entering the instrument aperture. This assumption is used primarily when viewing a blackbody source that fills the instrument field-of-view or when viewing Earth scenes with the emphasis on studying their spectral rather

than spatial variations. This assumption allows energy bundles to be diffusely emitted from the entire instrument aperture into the instrument.

The spectral distribution factor between the aperture and a spatial bin j on the detector surface for a spatially homogenous scene is defined as the fraction of energy entering the instrument aperture from the incident scene in wavelength band k of width $\Delta\lambda_k$ that is ultimately absorbed by detector surface element j through direct radiation and all reflections. This distribution factor can be estimated from the ray-trace by

$$D^k_{aperture-j} \cong \frac{N^k_{aperture-j}}{N^k_{aperture}} \tag{4.41}$$

where $N^k_{aperture-j}$ is the number of energy bundles absorbed by detector element j in wavelength band k and $N^k_{aperture}$ is the number of energy bundles emitted by the aperture in wavelength band k. The energy bundles are emitted from the aperture diffusely within a specified solid-angle represented by a cone angle. The zenith and azimuth angles, θ and ϕ, are calculated from

$$\theta = \sin^{-1}\left[\sin(\theta_{max})\sqrt{R_\theta}\right] \tag{4.42}$$

and

$$\phi = 2\pi R_\phi \tag{4.43}$$

where θ_{max} is the maximum zenith angle defining the cone angle, and R_θ and R_ϕ are random numbers uniformly distributed on the interval $(0, 1)$ [Walkup, 1996].

The power absorbed by any element j on the detector surface can be calculated from the total power entering the instrument aperture using the distribution factors. For a spatially homogeneous scene, the power incident to the instrument aperture in wavelength band k can be represented by

$$P^k_{aperture} = I^k \Omega A_{aperture} \tag{4.44}$$

where I^k is the band-averaged radiance in wavelength interval k, Ω is the instrument solid angle associated with the instrument's field of view, and $A_{aperture}$ is the area of the instrument aperture.

The power absorbed by a detector element j can be calculated by multiplying the power entering the aperture by the distribution factor between the aperture and element j for each wavelength band k, and then summing over all wavelengths. This idea can be expressed

$$P_{abs,j} = \sum_{k=1}^{N_\lambda} P^k_{abs,j} = \sum_{k=1}^{N_\lambda} P^k_{aperture} D^k_{aperture-j} \tag{4.45}$$

where $P^k_{abs,j}$ is the power absorbed by detector element j in wavelength band k, $P_{abs,j}$ is the total power absorbed by element j, and N_λ is the number of wavelength intervals.

4.1.6 Spatially inhomogenous scenes

A spatially inhomogenous field is one where the intensity incident at the instrument aperture varies across the scan and cross-scan angles. To accurately model this, it is necessary to subdivide the incident scene into discrete angular bins, with each bin representing a particular scan and cross-scan angle. Each angular bin measures $\Delta\theta$ by $\Delta\phi$, where θ represents the scan angle and ϕ represents the cross-scan angle.

The spectral distribution factor between the aperture and a spatial bin j on the detector surface for a spatially inhomogenous scene is defined as the fraction of energy entering the instrument aperture from angular bin (θ, ϕ) in wavelength band k that is ulimately absorbed by detector surface element j through direct radiation and all reflections. This distribution factor can be estimated from the ray trace by

$$D^k_{\theta,\phi-j} \cong \frac{N^k_{\theta,\phi-j}}{N^k_{\theta,\phi}} \ , \tag{4.46}$$

where $N^k_{\theta,\phi-j}$ is the number of rays absorbed by detector element j in wavelength band k, and $N^k_{\theta,\phi}$ is the number of rays emitted by the aperture in angular bin (θ, ϕ) and wavelength

band k.

For a spatially inhomogenous scene, the power incident to the instrument aperture in wavelength band k from angular bin (θ, ϕ) can be computed as

$$P^k_{aperture,\theta,\phi} = I^k_{\theta,\phi}\Omega_{\theta,\phi}A_{aperture} \tag{4.47}$$

where $I^k_{\theta,\phi}$ is the band-averaged radiance in wavelength interval k from angular bin (θ, ϕ), $\Omega_{\theta,\phi}$ is the solid angle associated with angular bin (θ, ϕ), and $A_{aperture}$ is the area of the instrument aperture. Haeffelin et al. [1997] showed that, for sufficiently small values of θ and ϕ, the solid angle for an angular bin can be well approximated by

$$\Omega_{\theta,\phi} \cong \Delta\theta\Delta\phi \tag{4.48}$$

The power absorbed by a detector element j can be calculated by multiplying the power entering the aperture by the distribution factor between the aperture and element j from each angular bin (θ, ϕ) for each wavelength band k, and then summing over all scan angles, cross-scan angles, and wavelengths. This idea can be expressed

$$P_{abs,j} = \sum_{k=1}^{N_\lambda} P^k_{abs,j} = \sum_{k=1}^{N_\lambda}\sum_{n_\theta=1}^{N_\theta}\sum_{n_\phi=1}^{N_\phi} P^k_{aperture,\theta,\phi}D^k_{\theta,\phi-j} \tag{4.49}$$

where $P^k_{abs,j}$ is the power absorbed by detector element j in wavelength band k, $P_{abs,j}$ is the total power absorbed by element j, N_λ is the number of wavelength intervals, N_θ is the number of subdivisions in the scan direction, and N_ϕ is the number of subdivisions in the cross-scan direction.

4.2 Dynamic electrothermal model

The radiometric modeling environment allows the user to model a generic bolometer detector assembly. The environment uses a finite difference electro-thermal model of a detector

assembly based on the thermistor bolometers used in the CERES instrument. Each detector consists of a thin sheet of Kapton bonded to an aluminum substrate by a thin layer of epoxy. A thermistor layer is then bonded to the Kapton with another thin layer of epoxy. Gold pads are deposited on each end of the thermistor, where platinum leads are attached. Another layer of epoxy and a layer of varnish are then deposited on top of the thermistor and gold pads. Finally a thin layer of black paint is deposited above the varnish. A schematic diagram of a single detector is presented in Figure 4.3.

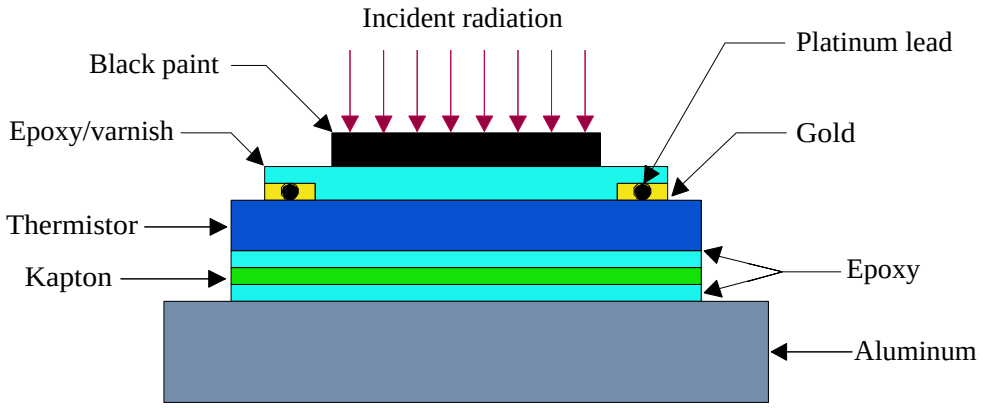

Figure 4.3: Schematic diagram of a CERES thermistor bolometer detector

The detector assembly is created by attaching two nearly identical detectors to their respective aluminum substrates. The two aluminum substrates are then clamped together back to back with an intervening layer of indium, resulting in a single structure with a detector on each side facing opposite directions. One detector is exposed to incident radiation and acts as the active detector, while the other detector is maintained in a nearly isothermal enclosure and acts as the reference, or compensating, detector. A schematic diagram of the entire detector assembly is presented in Figure 4.4.

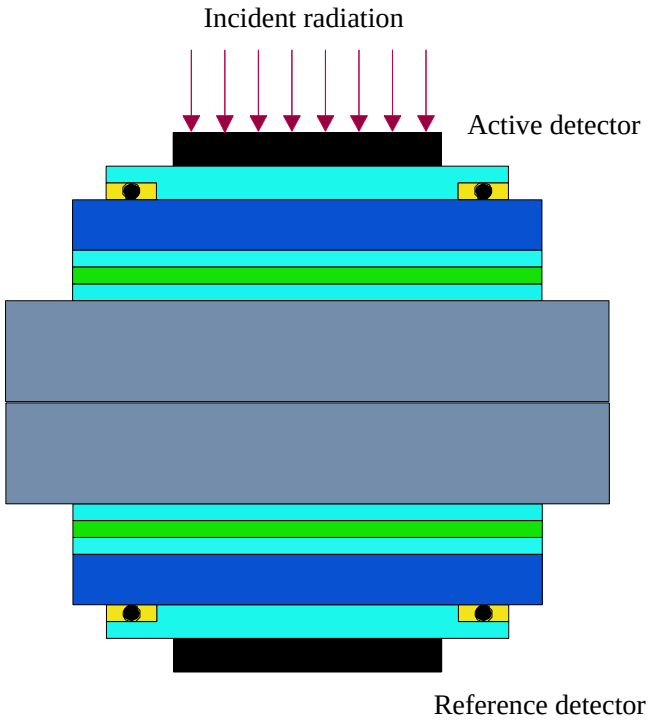

Figure 4.4: Schematic diagram of the CERES detector assembly

Although the detector model developed in this dissertation is based on the CERES design, it can be modified to model various bolometers by altering its physical properties and spatial dimensions. The thermal properties and the spatial dimensions are defined by two input files. The model can therefore be used to study the effect of different thermal impedance materials, thermistor materials, and absorber layers, as well as various detector sizes and thicknesses.

A three-dimensional transient finite-difference thermal diffusion model was developed to predict the change in temperature distribution in the active and reference detectors as well as in the substrate due to a spatially and temporally varying radiant input. Two-dimensional finite-difference electrical diffusion models of each thermistor layer were also developed to

predict the change in resistance in these layers due to changes in temperature. These models are coupled with the thermal model and use the temperature distributions in the layers to calculate temperature dpendent electrical resistivities in the electric models, which then provide the value of Joulean heating to the thermal model. The electric models are steady-state models calculated at each discrete time step of the unsteady thermal model.

At each discrete time step, the temperature distribution throughout the entire detector assembly is calculated with the thermal model. The resistivity of each control volume in the thermistor layers is calculated using the temperature distributions of each thermistor. This resistivity and the applied bias voltage across the detector are used to compute a discrete electric field distribution using the finite-difference method. Once the electric field is established, Ohm's law is applied locally to the control volumes along the borders where the voltage is applied to calculate the electric current entering and leaving the thermistor. This current is used with the bias voltage to determine the average resistance of each thermistor layer. These average resistances are used in the Wheatstone bridge circuit to calculate the new bridge voltage to be used in the next time-step.

At each discrete time-step, the bridge voltage is calculated and then used as the input to a model of the low-noise pre-amplifier and low-pass filter. The output of the filter is then converted to digital counts using 12-bit analog-to-digital conversion.

4.3 Thermal-electrical finite difference model

A finite difference model of each of the two thermistor layers was created to calculate the electric field distribution between the platinum leads at each discrete time step. A two-dimensional model is used, as opposed to a three-dimensional thermal model, because the thermistor layer is assumed to be electrically insulated by the layers of epoxy above and below in the z direction, and so current can flow only in the x and y directions. The model is steady-state and the electric field is recalculated at each time-step as the boundary

conditions change due to the transient thermal model.

The temperature distribution in the thermistor assembly is governed by

$$\frac{\partial}{\partial x}\left(k_x\frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial y}\left(k_y\frac{\partial T}{\partial y}\right)+\frac{\partial}{\partial z}\left(k_z\frac{\partial T}{\partial z}\right)+\dot{q}=\rho C_p\frac{\partial T}{\partial t} \tag{4.50}$$

where k is the thermal conductivity in W/m·K, T is temperature in K, $\dot{q}$ is the volumetric heat generation in W/m^3, ρC_p is the heat capacity per unit volume in J/m^3· K, and t is time in s. The thermal boundary conditions are:

- a combined incident heat flux and radiative boundary condition at the active detector surface
- a radiative boundary condition at all other exposed detector and substrate surfaces
- an adiabatic boundary condition along the vertical sides of the substrate.

The boundary conditions are illustrated in Figure 4.5. Equation 4.50 with the associated boundary conditions is solved using a fully implicit finite difference formulation.

The electric field in the thermistor layer is calculated based on the principle of conservation of electric charge, according to

$$\frac{\partial J_x}{\partial x}+\frac{\partial J_y}{\partial y}=0 \tag{4.51}$$

where J_x and J_y (A/m^2) are the current densities in the x and y directions, respectively. The current densities are related to the electrical conductivity, σ_e, $(\Omega\cdot\mathrm{m})^{-1}$ and the electric potential, E (V), according to

$$J_x=-\sigma_e\frac{\partial E}{\partial x}\qquad\text{and}\qquad J_y=-\sigma_e\frac{\partial E}{\partial y} \tag{4.52}$$

Equations 4.51 and 4.52 can be combined to obtain

$$\frac{\partial}{\partial x}\left(\sigma_e\frac{\partial E}{\partial x}\right)+\frac{\partial}{\partial y}\left(\sigma_e\frac{\partial E}{\partial y}\right)=0 \tag{4.53}$$

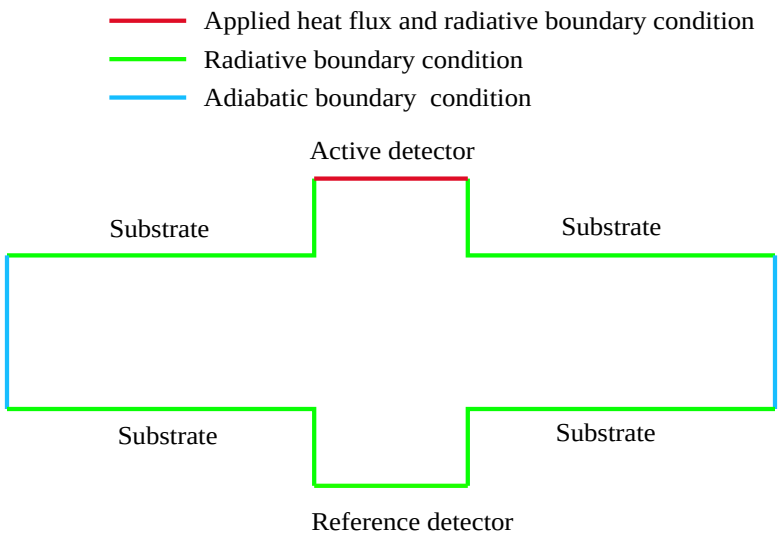

Figure 4.5: Thermal boundary conditions for thermistor bolometer finite difference model

which can be solved using the finite-difference method to provide the electric potential field $E = E(x, y)$.

The two-dimensional model is discretized into surface elements with the same x-y dimensions as the three-dimensional thermal model. The electrical resistivity of each element $\rho_e = 1/\sigma_e$ at temperature T is determined by

$$\rho_e = \rho_{e,o} \exp\left\{B\left(\frac{1}{T} - \frac{1}{T_o}\right)\right\} \tag{4.54}$$

where $\rho_{e,o}$ is the resistivity (Ω·m) of the thermistor layer element at a reference temperature T_o, and B (K) is a physical property of the thermistor material [Astheimer, 1984]. The electrical conductivity, σ_e, of an individual element can then be calculated from

$$\sigma_e = \frac{1}{\rho_e} = \frac{1}{\rho_{e,o} \exp\left\{B\left(\frac{1}{T} - \frac{1}{T_o}\right)\right\}} \tag{4.55}$$

At each discrete time-step, the temperature distribution in the thermistor layer is determined

from the thermal model. The temperature, T, of each element in the thermistor is then used to calculate the electrical conductivity of each element according to Equation 4.55. The finite difference coefficients are then calculated for each element and the resulting set of equations is solved for the electric potential field, E. The electric potential at the thermistor edges under the platinum leads is supplied by the bias voltage calculated in the electronic circuit model, and the other edges are assumed insulated.

Once the electric potential field is known, the current through the thermistor can be calculated by applying Ohm's law locally at either end containing the platinum leads, according to

$$I = \sigma A \frac{E_i - E_{i-1}}{\Delta x} \tag{4.56}$$

Once the current across the thermistor is known, an equivalent resistance R_{equiv} for the thermistor can be calculated from

$$R_{equiv} = \frac{\Delta E_{thermistor}}{I} \tag{4.57}$$

The equivalent resistance is the used in the electronic circuit model to calculate the overall instrument output.

4.4 Electronic signal conditioning model

An analytic model of a typical pre-amplifier circuit, based on the CERES pre-amplifier system, was adapted from a previous model developed by Priestley [1997]. The model consists of an amplifier and low-pass filter, and assumes ideal electronic components. A detailed diagram of the circuit is displayed in Figure 4.6.

The output of the low-pass filter may be represented as

$$V_{out} = -\frac{1}{R_{12}C1} \int \left[\frac{R_{11}}{R_8} \frac{R_5 + 2R_6}{R_5} \right] \tag{4.58}$$

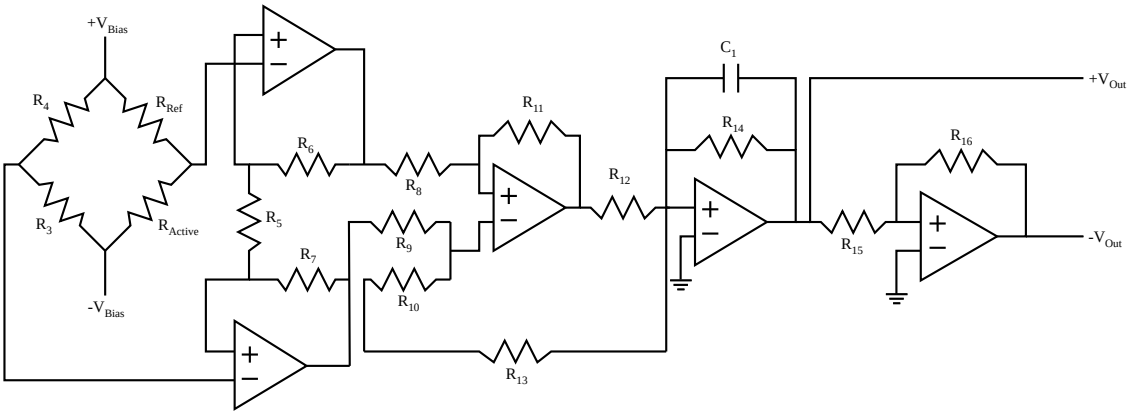

Figure 4.6: Schematic diagram of pre-amplifier circuit

where the bridge output is given by

$$V_{bridge} = V_2 - V_1 \tag{4.59}$$

For frequencies sufficiently below

$$f = \frac{1}{2R_{14}C_1} \tag{4.60}$$

the low-pass filter acts primarily as an amplifier [Priestley, 1997] with a gain described by

$$V_{out} = \frac{R_{14}}{R_{12}} \left[\frac{R_{11}}{R_8} \frac{R_5 + 2R_6}{R_5} \right] \tag{4.61}$$

After leaving the pre-amplifier, the signal is converted to a digital output using a 12-bit analog-to-digital converter. This provides 4096 bits to cover a full-scale voltage range of 0-to-10 V and a count conversion factor of 409.5 V/count.

Chapter 5

End-to-end model of the CERES instrument

5.1 CERES Instrument

The CERES instrument contains three scanning thermistor bolometer assemblies: a visible channel that measures earth-reflected solar radiance in the 0.3-5.0 μm range, a total channel that measures both earth-reflected solar radiance and earth-emitted radiance in the 0.3-100 μm range, and an atmospheric "window" channel that measures the portion of earth-emitted radiance lying in the 8-12 μm range. Each sensor assembly consists of a telescope containing a Cassegrain optical system; a detector assembly consisting of two bolometers, an active and a reference detector mounted on opposing sides of an aluminum substrate; and the associated signal-conditioning electronics.

The optical system consists of a forward baffle, an $f1.8$ Cassegrain optical module consisting of an 18-mm diameter silvered spherical primary mirror and an 8-mm diameter silvered spherical secondary mirror. The shortwave and window channels have filters located before the "spider" (the three-legged structure that supports the secondary mirror) and in front

of the active bolometer detector. The shortwave channel filters consist of two 1-mm thick fused, waterless quartz elements, while the window channel filter system consists of a 1-mm thick zinc sulfide filter and a 0.5-mm thick cadmium telluride filter. The incident radiance is collected by the optical system and projected onto a 0.75-by-1.50-mm truncated diamond-shaped precision aperture mounted directly in front of the active bolometer. The aperture restricts the sensor to a 1.3-by-2.6-deg field-of-view. A schematic diagram of all participating radiative surfaces and their geometries is given in Figure 5.1.

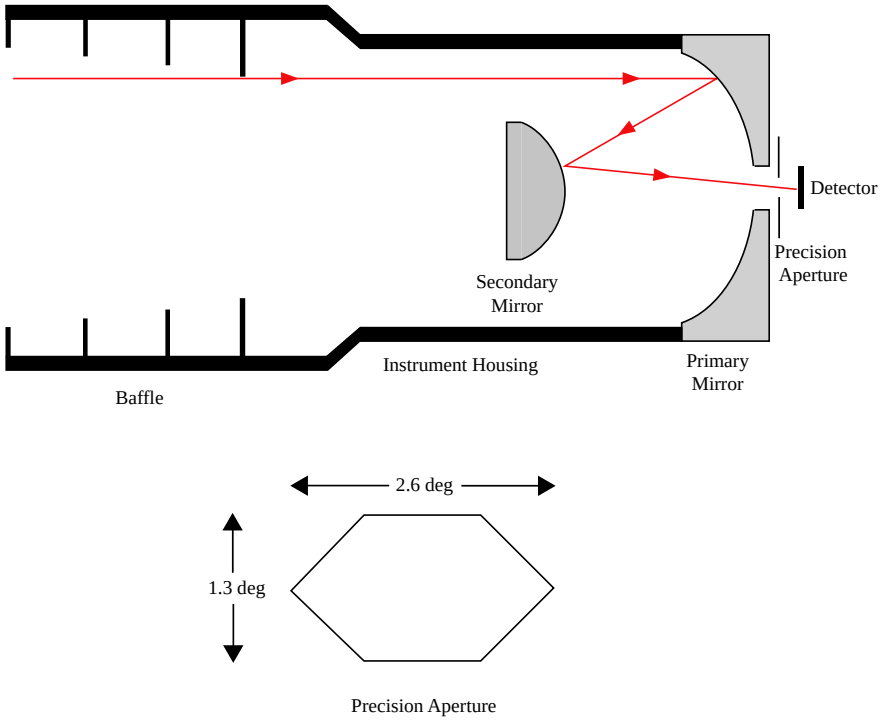

Figure 5.1: Schematic diagram of CERES optical system

The detector assembly consists of an active bolometer and a reference bolometer mounted on opposite sides of an aluminum substrate. Each bolometer is a rectangle measuring approximately 1.6 by 3.0 mm with a 1.6-by-1.6-mm absorber area and is approximately 40 μm thick. They are mounted on separate aluminum heat-sink disks that are 30.76 mm in diameter and 3.86 mm thick, and are joined via a 100-μm-thick layer of sputtered indium. The heat sinks

are maintained at a constant temperature of 38°C using actively controlled 2.3-W electric heaters. The two detectors are electrically connected as two adjacent legs in a Wheatstone bridge circuit. As the radiant energy is collected and focused onto the active bolometer's absorber paint layer, the absorbed energy is converted to heat and causes a temperature rise in the active bolometer, which causes a change in the electrical resistance of the thermistor layer of the active bolometer. This change in resistance is governed by

$$\rho = \rho_o \exp\left[B\left(\frac{1}{T} - \frac{1}{T_o}\right)\right] \tag{5.1}$$

where ρ_o is the electrical resistivity at a reference temperature T_o and B is a material property of the thermistor. The change in resistance causes a deflection of the bridge voltage and the output signal is passed through the low-noise pre-amplifier and low-pass filter before being sampled, as described in Chapter 4.

5.2 End-to-end numerical model of CERES scanning radiometer

A complete end-to-end numerical model of a single channel of the CERES instrument has been developed, integrating an optical Monte Carlo ray-trace (MCRT) model, a three-dimensional transient electro-thermal finite-difference diffusion model, and a model of the signal-conditioning electronics. This model was created as an extension of previous modeling efforts by Bongiovi [1993], Priestley [1997], Haeffelin [1997], and Coffey [1998]. It is intended to increase the computational efficiency and flexibility of these previous models in order to be integrated with other numeric models.

5.3 Optics results

This section outlines the results of the MCRT model of the CERES optical system developed in the radiometric instrument modeling environment. The model was created by specifying the two-mirror system and appropriate parameters that correspond to the geometry and dimensions of the as-built CERES optical system. The model was used to simulate radiative transfer in the instrument from the instrument aperture to the active detector surface for various conditions. A description of the spectral characteristics and performance of the model is also provided.

5.3.1 Description of the optical model

A MCRT model of the CERES optical system was developed based on nominal dimensions and properties of the as-built CERES radiometric channel. The optical system consists of a spherical primary mirror and a spherical secondary mirror in a Cassegrain configuration. The curvature, vertex radius, and mirror-slice depth that relate the CERES telescope to the

input parameters for the radiometric modeling environment shown in Figure 4.2 are given in Table 5.1.

Table 5.1: Mirror dimensions for CERES telescope

	Depth c (mm)	Mirror slice width a (mm)	Location of vertex, Z_v (mm)
Primary mirror	1.25	9.50	73.38
Secondary mirror	0.25	4.11	62.43

The instrument housing, baffle, mirror supports, and precision aperture were modeled using quadric surfaces according to nominal CERES dimensions. The absorptivity of the mirrors and detector are wavelength dependent, and were discretized into wavelength bands. Representative reflectivities are plotted as a function of wavelength for the primary and secondary mirrors in Figure 5.2. The absorptivity of all other surfaces were assumed to be gray at a value of 0.9 and the specularity ratio was assumed to be gray for all surfaces at a value of 1.0. The modeled and nominal instrument spectral response (optical throughput) values are shown in Figure 5.3.

5.3.2 Convergence of the MCRT model

A study was performed to determine the accuracy of the MCRT model as a function of the number of energy bundles emitted from the emitting surfaces and to establish the appropriate number of energy bundles to be fired for the desired analysis. Two types of emission are considered in the optical systems model: collimated radiation from a specified scan and cross-scan angle, and diffuse radiation within a specified solid-angle. A convergence study for both types of emission is necessary, as the ideal number of energy bundles emitted for one case may not necessarily correspond to that for the other case.

For the case of collimated radiation, energy bundles were fired parallel to the optical axis, corresponding to scan and cross-scan angles of 0 deg. The distribution factor from the

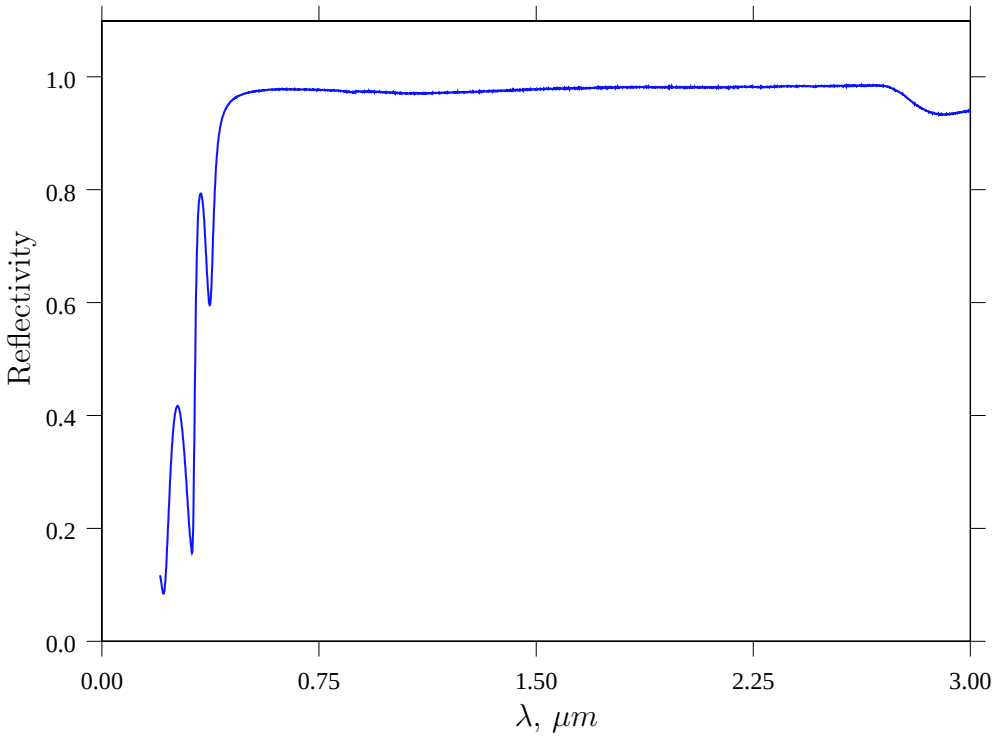

Figure 5.2: Nominal reflectivity values for primary and secondary mirrors used in CERES end-to-end model

aperture to the detector for each case corresponding to a specific number of emitted energy bundles is summarized in Table 5.2 and is displayed in Figure 5.4.

The values for the distribution factor converge to a value of 0.327 for a number of emitted energy bundles equal to one million. However, it should be noted that the difference between the values of the distribution factor for the case of ten thousand rays and one million rays is on the order of half a percent, which indicates that relatively accurate values for the distribution factor can be obtained with a fairly small number of rays, for this particular application.

For the case of diffuse radiation in a specified solid angle, energy bundles were fired diffusely

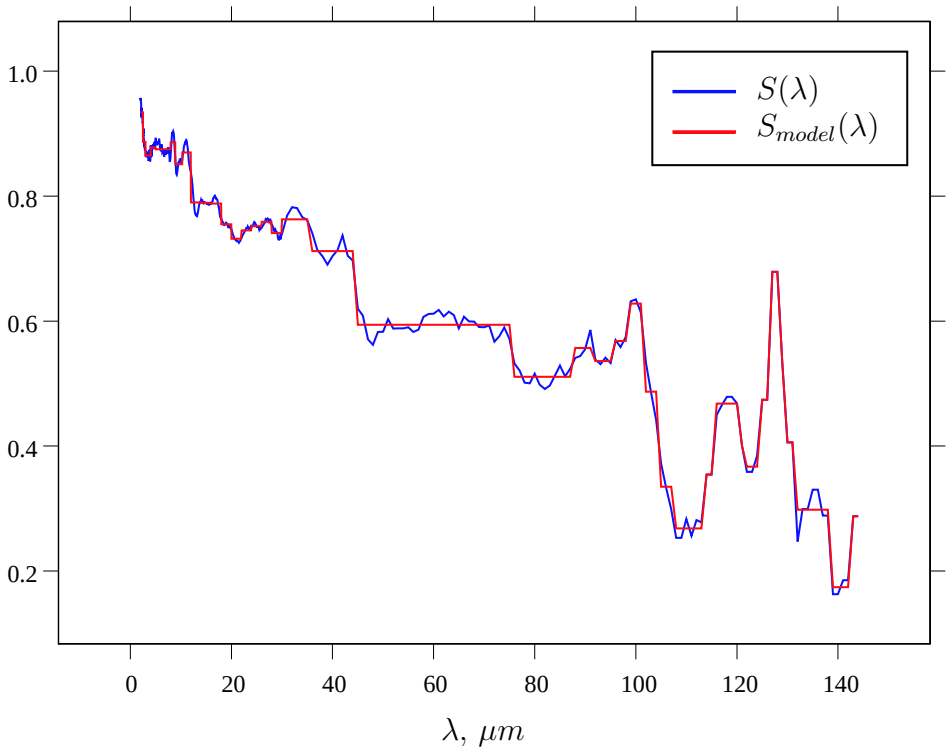

Figure 5.3: Measured and modeled nominal spectral response (optical throughput)

within a cone angle of 1.6 deg corresponding to a solid angle of 2.45×10^{-3} sr. The distribution factor between the aperture and detector for each case corresponding to a specific number of energy bundles is summarized in Table 5.3 and is displayed in Figure 5.5.

The values for the distribution factor converge to a value of 0.101 for a number of emitted energy bundles equal to five million. However, the value of the distribution factor convereges to within 2 percent of the final value with only five thousand energy bundles, again raising the possibility of fairly accurate values of the distribution factor with a relatively small number of energy bundles emitted.

Table 5.2: Distribution factor and percent difference for increasing number of energy bundles traced for the case of collimated incident radiation

Number of Rays	D'_{oj} (-)	Percent change (%)
1000	0.315	-
5000	0.323	+2.48
10000	0.325	+0.62
50000	0.325	0.00
100000	0.324	-0.31
500000	0.327	+0.92
1000000	0.327	0.00
5000000	0.327	0.00
10000000	0.327	0.00

5.3.3 Solid angle convergence study

For the case of diffuse emission within a specified solid angle, it is necessary to determine the solid angle associated with the instrument's field of view. To accomplish this, it is necessary to find the minimum cone angle, θ, that encompasses the entire field-of-view. This can be done by varying θ and calculating the absorbed power within the associated solid angle for a uniform radiant source at the instrument aperture. When the absorbed power is no longer increasing as a function of the cone angle, then the full field of view is encompassed. A ray trace using 10,000 energy bundles was performed for several cases corresponding to increasing cone angles, and the resulting distribution factors, total incident power, and total absorbed power within the solid angle are displayed in Table 5.4 and in Figure 5.6.

The values for absorbed power on the detector flake converge to approximately 12.3 μW at a cone angle of 1.3 deg. This corresponds to the theoretical CERES field of view of 1.3 deg by 2.6 deg. The reason that the absorbed power fluctuates about 12.3 μW in the numerical study is that, as the cone angle, and thus the solid angle increases, a fixed number of energy bundles is being fired into an increasingly larger solid angle, and thus the results have a decreasing accuracy as the solid angle increases, causing some fluctuation in the values for

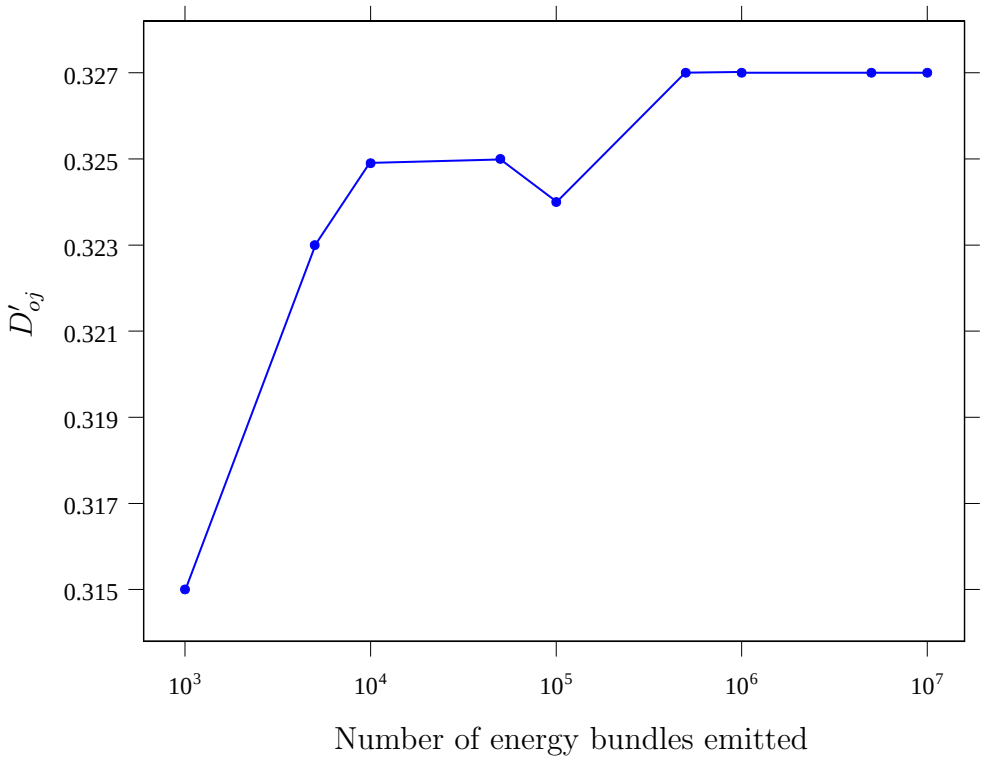

Figure 5.4: Distribution factor values for increasing number of emitted collimated energy bundles

absorbed power as θ increases. For all further simulations of the CERES optical system using diffuse emission, a cone angle of 1.6 deg is used in the ray trace.

5.3.4 On-axis collimated radiation

Collimated radiation entering the instrument aperture is simulated by emitting energy bundles into the instrument from the aperture parallel to the optical axis. This is useful for evaluating properties such as optical throughput and the size of the blur circle on the detector surface. Figure 5.7 is a plot of the spatial coordinates in the x and y directions of the

Table 5.3: Distribution factor and percent difference for increasing number of energy bundles for the case of diffuse incident radiation within a specified solid angle

Number of energy bundles	D'_{oj} (-)	Percent change
1000	0.106	-
5000	0.103	-2.83
10000	0.102	-0.98
50000	0.103	+0.97
100000	0.103	0.00
500000	0.102	-0.98
1000000	0.102	0.00
5000000	0.101	-0.99
10000000	0.101	0.00

energy bundles absorbed on the detector surface. For this case, ten thousand energy bundles were emitted on-axis from the instrument aperture. We can see that the radius of the blur circle is approximately 0.07 mm, which compares well with values published by Bongiovi [1993], Haeffelin [1997], and Coffey [1998]. It is also interesting to note that the image of the three-legged secondary mirror support, or "spider," is visible on the detector surface. The optical throughput, defined in this dissertation as the fraction of energy entering the aperture that is ultimately absorbed by the detector, is 0.32 for this case.

5.3.5 Diffusely emitted radiation within a specified solid angle

The more common case of diffuse radiation entering the instrument aperture is modeled by emitting energy bundles diffusely from the aperture surface. While traditionally diffuse energy bundles are emitted in a hemisphere, the actual instrument has a limited field of view and cannot "see" incident energy outside this field of view, which has a value of approximately 1.3 deg from nadir. Emitting diffusely into a hemisphere therefore results in a large number of energy bundles being emitted that cannot reach the detector surface. A more efficient

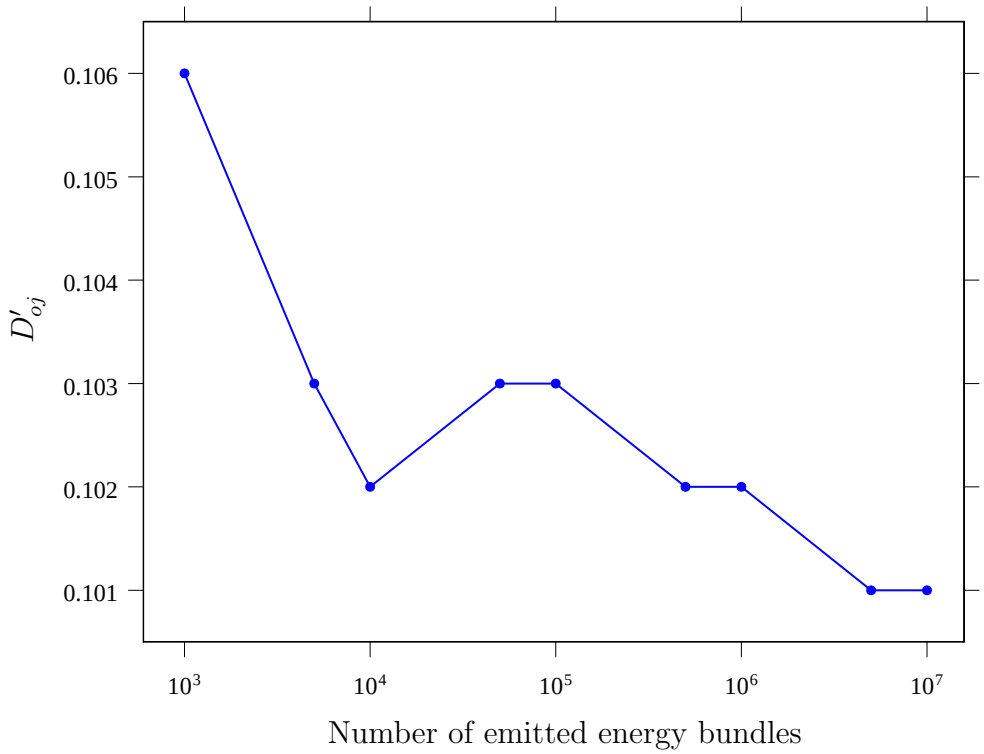

Figure 5.5: Distribution factor values for increasing number of diffusely emitted energy bundles

approach, suggested by Walkup [1996], is to emit energy bundles diffusely into a limited field of view, which can be defined as a cone defined by a cone angle θ. Figure 5.8 is a plot in the x and y directions of the energy bundles absorbed on the detector surface. For this case, ten thousand energy bundles were emitted from the instrument aperture inside a solid angle specified by a cone angle of 1.6 deg. Unlike the case of collimated radiation, which results in a finite blur circle on the detector surface that is much smaller than the precision aperture, the spatial distribution of absorbed energy on the detector surface is defined by the precision aperture, as can be seen in Figure 5.8. Also, the "spider" is not readily visible in the absorption pattern.

Table 5.4: Values for the distribution factor, total incident power, and absorbed power within a specified solid angle for an increasing cone angle

θ	D_{oj}' (-)	Incident power (μW)	Absorbed power (μW)
0.1	0.325	0.476	0.154
0.3	0.324	4.28	1.39
0.5	0.324	11.9	3.85
0.7	0.310	23.3	7.23
0.9	0.267	38.5	10.3
1.1	0.205	57.6	11.8
1.3	0.152	80.4	12.3
1.5	0.116	107.0	12.4
1.7	0.090	137.5	12.3
1.9	0.072	171.7	12.3
2.1	0.058	209.8	12.2
2.3	0.049	297.3	12.3

5.3.6 Optical point spread function

The optical point spread function (OPSF) of the instrument is the angular optical response of the instrument. This can be represented as the fraction of energy incident from a particular direction that is absorbed by the detector, and is a function of the angular coordinates that define the direction of incident radiation. A theoretical value of the OPSF can be calculated from the MCRT instrument model by emitting collimated energy bundles from a specific scan and cross-scan angle combination and recording the fraction of energy absorbed. This is repeated for all possible combinations of scan and cross-scan angles within the range of interest, which in this case is the instrument field of view. This process was performed using the current CERES model, and the resulting OPSF is shown as a contour plot in Figure 5.9 and as a surface plot in Figure 5.10. In both plots the OPSF has been normalized to a value of unity and is plotted against the scan and cross-scan angles. The OPSF is nearly uniform across the instrument field of view and then falls off sharply at approximately 1.2 deg in the cross-scan direction and 0.5 deg in the scan direction, eventually going to a value of zero at

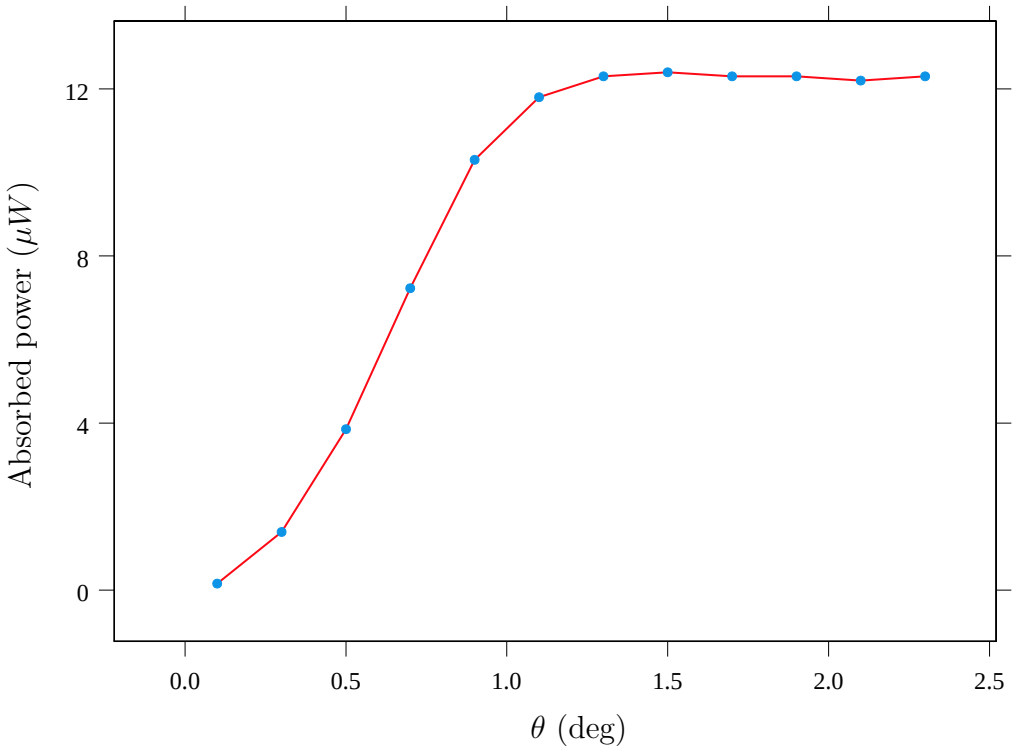

Figure 5.6: Absorbed power for increasing cone angle

1.3 deg in the cross-scan direction and 0.65 deg in the scan direction. The OPSF does not have a true "top-hat" response but instead a gradual drop-off due to the finite size of the blur circle, and the roll-off represents the angular region where portions of the blur circle are actually focused on the precision aperture mask. The shape of the OPSF is governed by the shape of precision aperture, with the diamond shape clearly visible in both plots.

5.3.7 Instrument point spread function

The instrument point spread function represents the dynamic angular response of the instrument, and incorporates both optical and electrothermal effects. It can be measured by

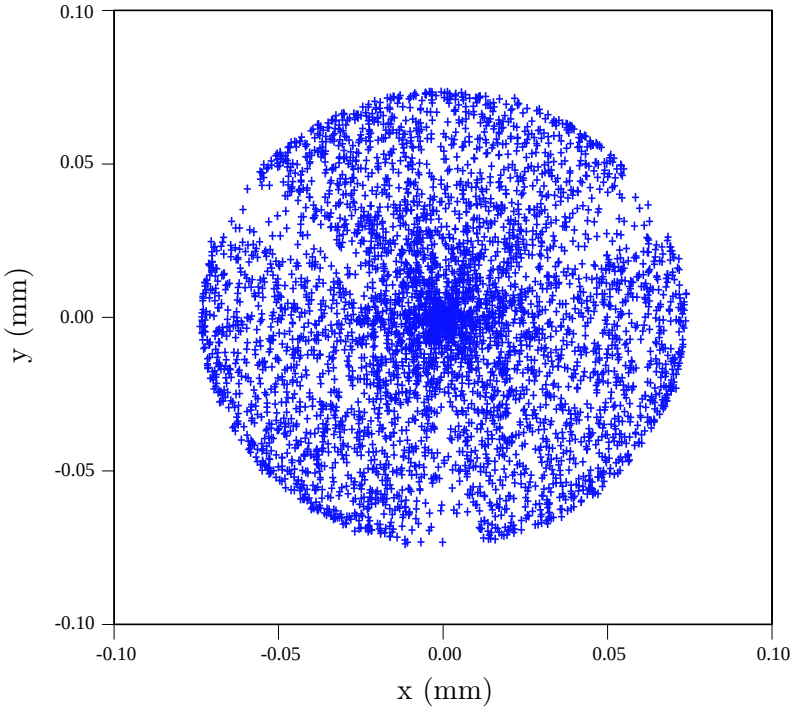

Figure 5.7: Scattergram for ray trace with collimated incident radiation

scanning across a point source at a specified cross-scan angle and measuring the resulting instrument output as a function of time, which can be converted to a function of scan angle with knowledge of the scan rate (63.5 deg/s in this case). This is repeated for all cross-scan angles of interest, and the resulting point spread function can be presented as a function of scan and cross-scan angles. This process was simulted with the instrument MCRT model by simulating an Earth-scan with the instrument model. Incident radiation is emitted by the aperture at a fixed cross-scan angle and at a scan-angle corresponding to the current time step. The direction of incident radiation and thus total power absorbed by the detector is updated at each time step. The absorbed power was used as the boundary condition to the electrothermal finite-difference model, and the boundary conditions were updated at each time step and the resulting instrument output was recorded at each time step. The resulting point spread function is shown as a contour plot in Figure 5.11 and as a surface plot in Figure 5.12. Unlike the OPSF, the PSF does not have a flat peak but has a very narrow

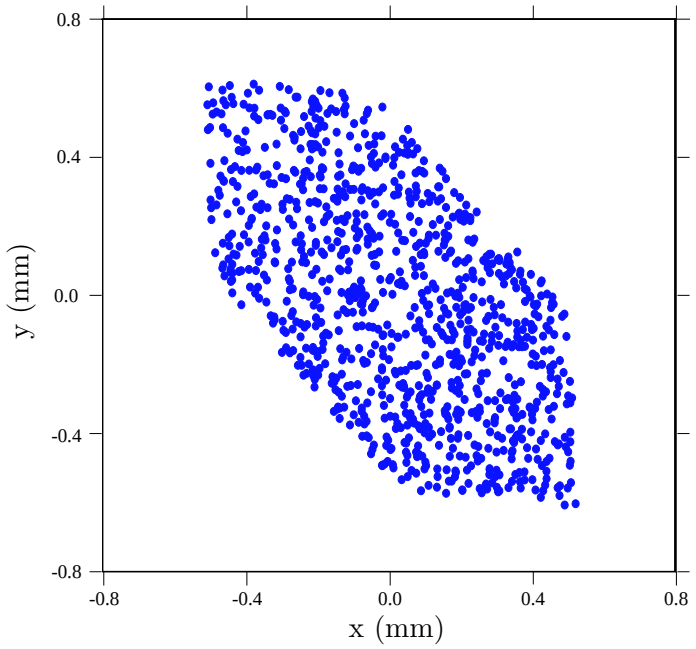

Figure 5.8: Scattergram for raytrace with diffuse incident radiation

peak at a scan-angle of approximately 0.5 deg instead of 0 deg. The peak is shifted due to the transient electrothermal time constant of the detector. The detector response also begins at approximately -0.7 deg and extends past 1.5 deg in the scan direction. This indicates that at any given instant, the detector signal is a result of the scene the instrument is viewing and the scene it has just viewed. The point spread function is an important analytic tool in post-processing the instrument output to account for these effects.

5.3.8 Absorbed energy on detector surface

The analyses in the preceding sections were all performed assuming gray surfaces. This was done to reduce the run-time of the MCRT models by only having to run a ray trace for a single wavelength band. This can be justified because the blur-circle is only a function of the types, locations, and specularity of the mirrors and other optical components, which are not functions of wavelength. Likewise, the shape of the OPSF would not vary with wavelength,

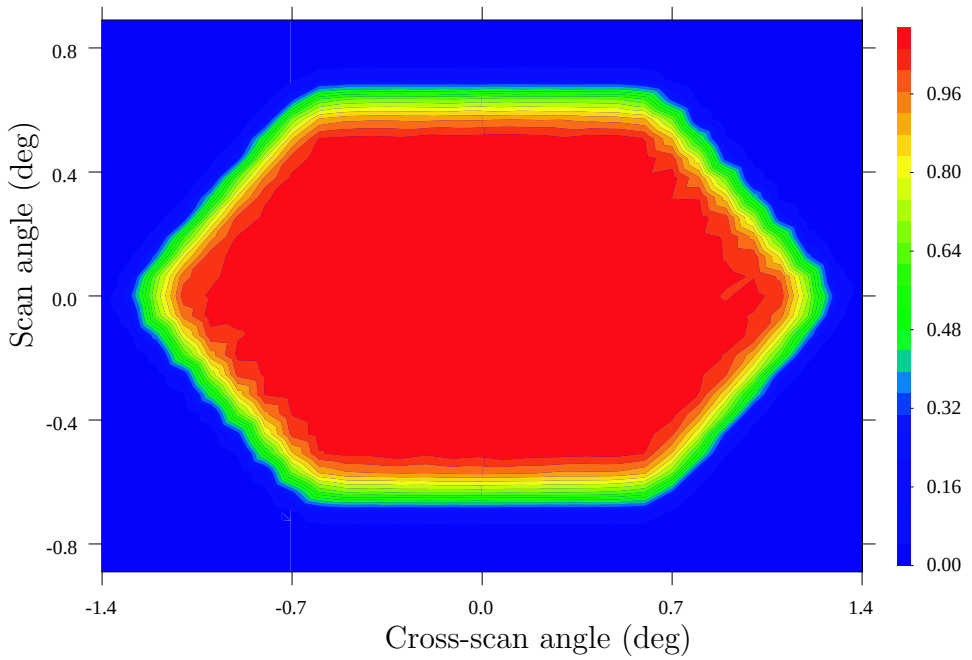

Figure 5.9: Optical point spread function contour plot

although the overall magnitude would be wavelength dependent. However, as we have only presented normalized values, these latter again would not vary with wavelength. The absorbed power on the detector surface, however, is dependent on not only the spectral nature of the optical system but on the emitting source spectral distribution as well. It is therefore necessary to discretize the optical properties into N_k discrete wavelength intervals, with k representing the kth wavelength interval of width $\Delta\lambda_k$. A ray trace was performed for each wavelength interval, and distribution factors between the aperture, or angular bin, and each discrete element j on the detector surface were calculated for each wavelength interval. The total energy entering the aperture in each wavelength interval was calculated by integrating the spectral radiance over that wavelength interval and multiplying by the solid angle and the area of the instrument aperture. The total power is then used with the distribution factor

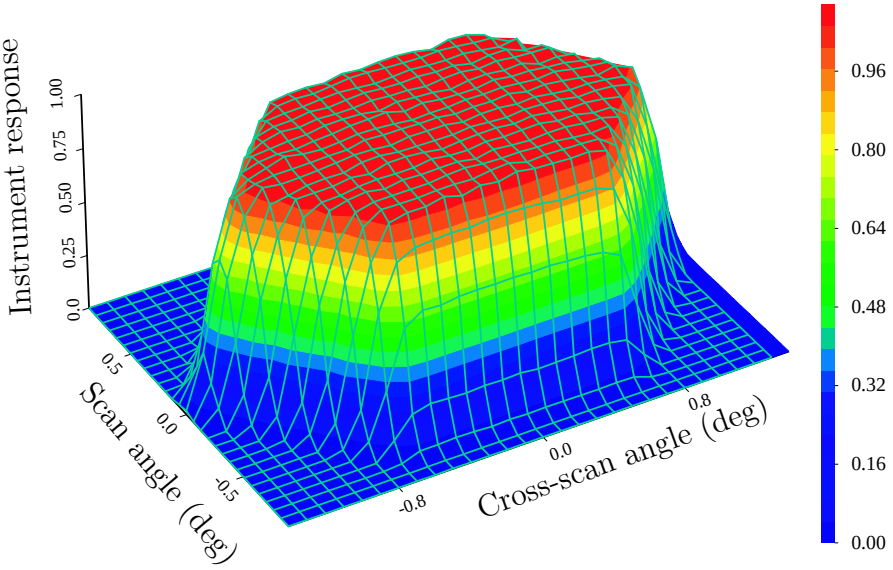

Figure 5.10: Optical point spread function surface plot

to find the absorbed power in each detector surface element in that particular wavelength interval. This is repeated for each wavelength interval, and the resulting absorbed powers are then summed over all wavelength intervals to find the total absorbed power in each surface element on the detector. This was done with the CERES radiometric model, using the spectral properties from Figure 5.2 and the incident radiation corresponding to a blackbody source at 359 K, or a radiance of 300 W/m^2-sr. The spatial distribution of the total absorbed power on the detector surface is shown in Figure 5.13. The spatial distribution is governed by the shape of the precision aperture, as can be clearly seen in Figure 5.13, and the absorbed power is essentially uniform over this area, as we would expect for a diffuse source with a spatially uniform radiance. Some minor variability exists in the absorbed power, which can be attributed to numerical sampling due to the finite number of energy bundles emitted in

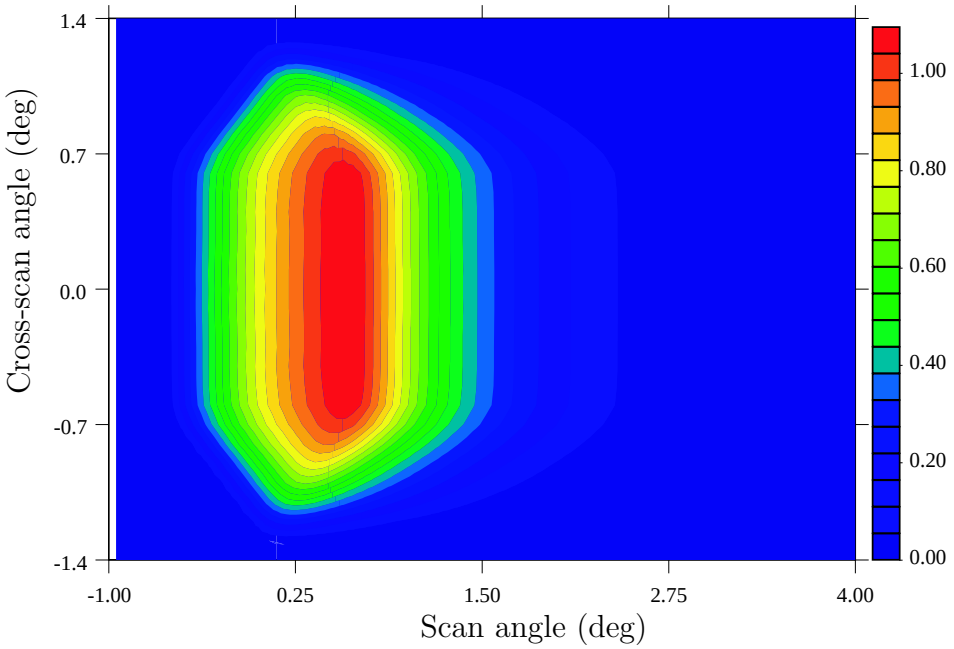

Figure 5.11: Instrument point spread function contour plot

the ray trace.

5.4 Detector assembly electrothermal model results

The electrothermal model of the detector assembly is a three-dimensional, transient finite difference thermal model of the entire detector assembly coupled with steady-state finite difference electrical models of both the active and reference thermistor layers. The boundary conditions for the thermal models are fixed with time on all surfaces except on the active detector surface, which has an applied heat flux boundary condition, received from the optical model, that can vary with time. The model calculates the overall temperature distribution in the detector assembly, the electric field in both thermistors, and the instrument output

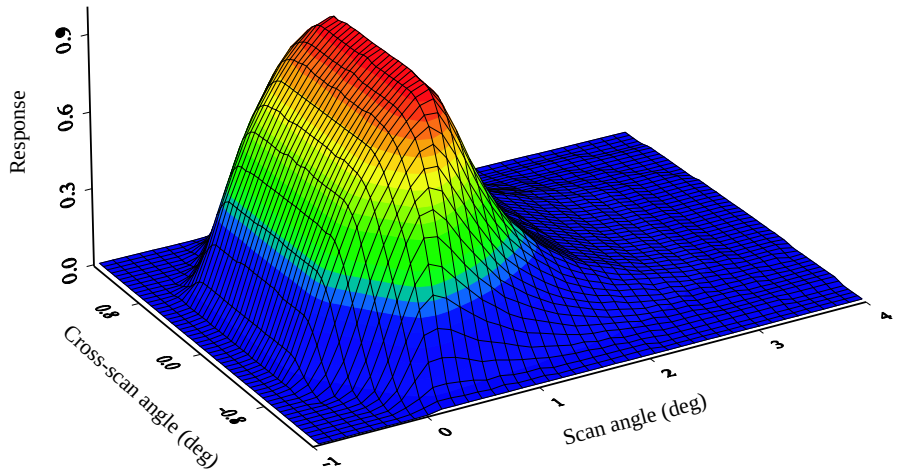

Figure 5.12: Instrument point spread function surface plot

voltage from the bridge circuit, amplifier circuit, and the final digital output for each time step. Various studies were performed with the electrothermal model and the results are outlined in the following sections.

5.4.1 Temperature distribution when exposed to a blackbody source

The CERES radiometric model was used to simulate the instrument viewing a blackbody source at 359 K, which produces a radiance of approximately 300 W/m^2. The incident radiation was treated as a step input, with the detector assembly in thermal equilibrium before being exposed to the step input. The electrothermal model was allowed to evolve to steady-state, with the instrument output recorded for each time step. A discrete time step of 0.001 s was used in the study. The detector surface was discretized into 16×16 elements in the x- and y-directions, the subsequent layers were discretized into 16×24 elements in

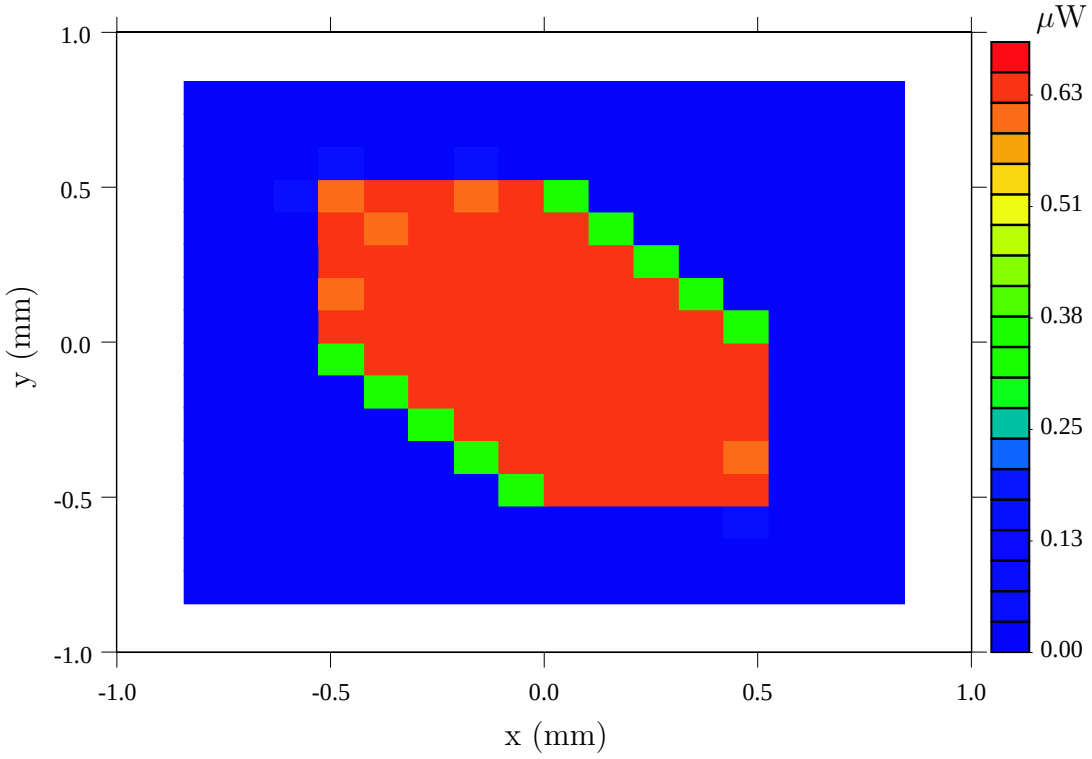

Figure 5.13: Absorbed power on active flake when viewing a blackbody source at 359K

the x- and y-directions, and the aluminum substrate was discretized into 24×32 elements in the x- and y-directions. All temperatures reported in the following studies are temperature differences referenced to the initial substrate temperature of 311.15 K. Nominal values for layer thickness, thermal conductivity, density, and specific heat were used. These nominal values are displyed in Table 5.5.

The steady-state temperature distribution on the active detector surface is shown in Figure 5.14. The temperature distribution reflects the shape of the precision aperture, with a peak temperature defect of 1.18 K over the area under the precision aperture, and diffusing out to a value of 1.16 K in the areas not directly exposed to incident radiation.

Table 5.5: Nominal values used in CERES electrothermal detector model

	z (μm)	k (W/m-K)	ρ (kg/m^3)	C_p (J/kg-K)
Epoxy (μm)	1.0	0.13	1200	1045
Kapton (μm)	7.52	0.12	1420	1091
Thermistor (μm)	15.0	8.36	5000	752.4
Absorber	10.6	0.209	1400	668.8
Varnish/epoxy	7.5	0.1	1150	1000
Gold	0.5	293	19,320	129.6
Aluminum	3668	237	2700	903
Indium	2.54	80.86	11,480	280.3

The steady-state temperature distribution on the reference detector surface is shown in Figure 5.15. The reference detector is not exposed to incident radiation from the scene, and so the temperature distribution on the reference detector surface is a result of the energy released in the thermistor layer due to self-heating leaving the detector by means of radiant emission from the surface. The maximum surface temperature defect of 1.19 K occurs at the center of the x-y plane of the detector, and the temperature defect decreases in a radial direction from the center to a minimum value of 1.18 K.

5.4.2 Thermistor layers

The steady-state temperature distribution in the active thermistor layer is shown in Figure 5.16. The shape of the precision aperture, while visible to some degree, is much less prominent than at the detector surface, with the temperature variation being more spatially uniform due to the heat diffusion throughout the layer. The peak temperature defect in the layer is 1.19 K, which is actually higher than at the surface due to the effects of the electrical self-heating in the thermistor layer. In other words, net heat is lost from the active flake during operation.

The steady-state temperature distribution in the reference thermistor layer is shown in Fig-

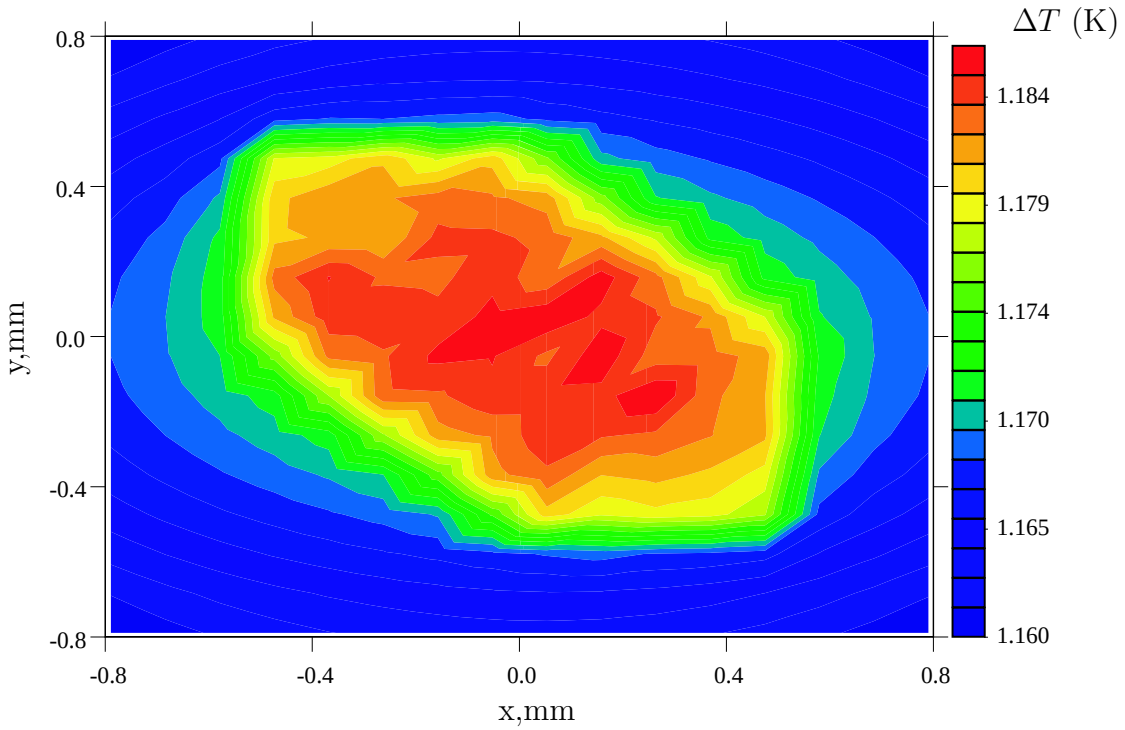

Figure 5.14: Steady-state temperature distribution on active flake when instrument is viewing a blackbody source at 359 K

ure 5.17. As the reference detector is not exposed to incident radiation, the active heat transfer mechanism in the reference detector is the balance among heat generation in the thermistor, radiation out of the top of the detector surface, and diffusion into the aluminum heat sink. The temperature distribution is elliptical, with a peak temperature defect of 1.192 K in the center and decreasing temperatures in both the x- and y-directions.

The steady-state values of the resistivity ρ (Ω-m) in the active thermistor layer when viewing a 359-K blackbody source are shown in Figure 5.18. Again the spatial distribution is influenced by the shape of the precision aperture, with the minimum value of 1.4811 Ω-m in the center of the layer and the values increasing in both the x- and y-directions, with

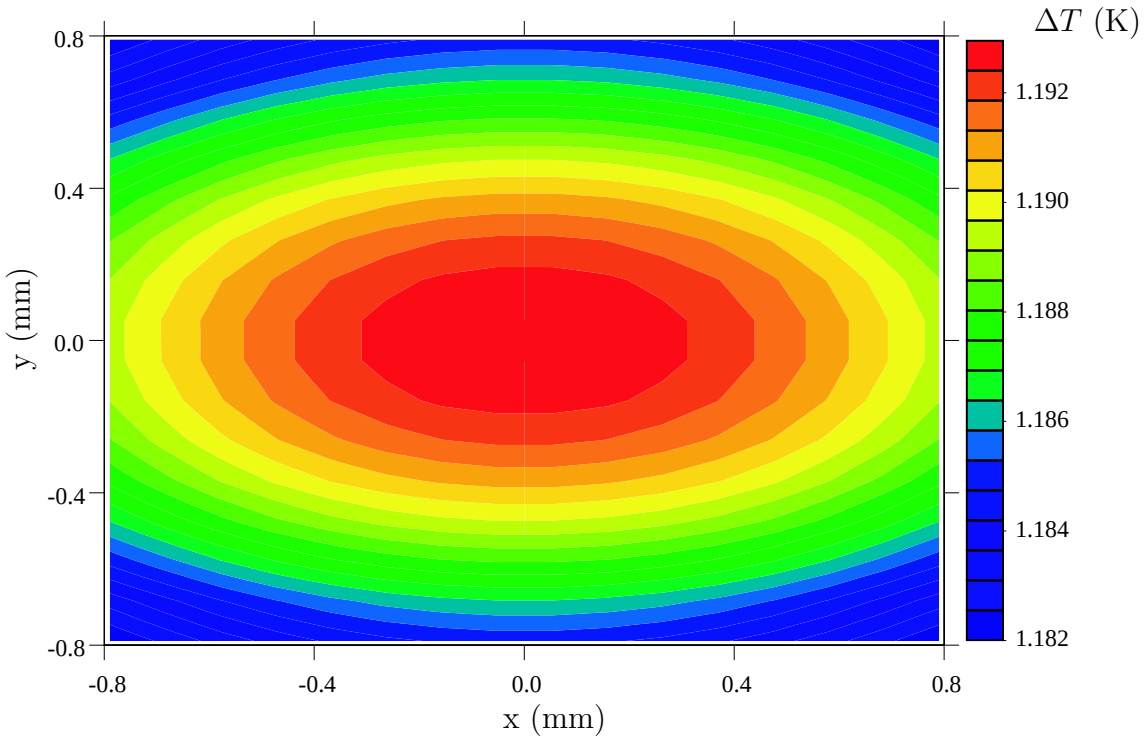

Figure 5.15: Steady-state temperature distribution on reference flake when instrument is viewing a blackbody source at 359 K

the maximum values at the four corners of the x-y-plane of the thermistor. The spatial distribution of resistivity mirrors that of the temperature distribution, with the minimum being at the center instead of the maximum, due to the fact that the resistivity is inversely related to the tempertature, as revealed in Equation 5.1.

The spatial distribution of resistivity for the reference thermistor layer is shown in Figure 5.19. The distribution is similar to that of the active layer, with a minimum value of 1.4809 Ω-m in the center and increasing values further from the center. Unlike the active thermistor, the distribution has a circular shape which mirrors that of the reference thermistor layer's temperature distribution.

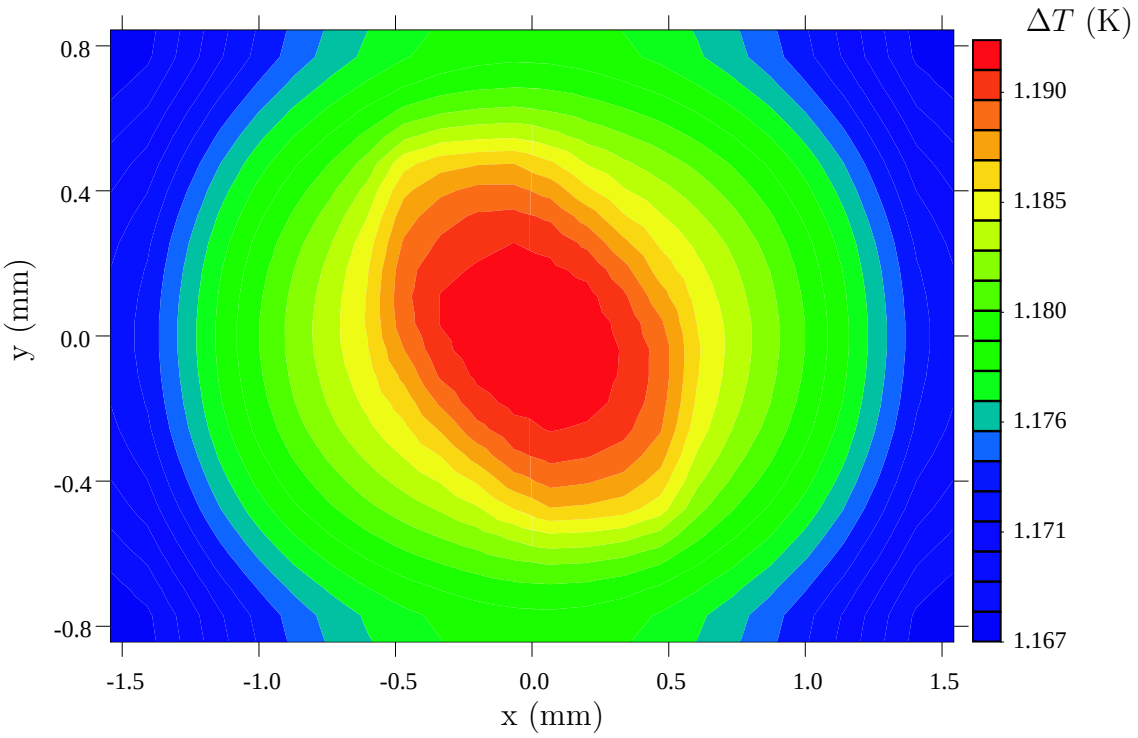

Figure 5.16: Steady-state temperature distribution in active thermistor layer when instrument is viewing a blackbody source at 359 K

5.4.3 Temperature distributions when instrument is viewing cold space

Another case of interest is when the CERES instrument is viewing cold space. The "space-look" is used as the reference with respect to which the signal from the earth-scan is measured. Cold space can be modeled as a blackbody at 4 K, and a simulation can be performed using the radiometric model in a similar manner to that described in the preceding sections. The temperature distribution on the detector surface is shown in Figure 5.20, and the temperature distribution in the active thermistor layer is shown in Figure 5.21. For both cases

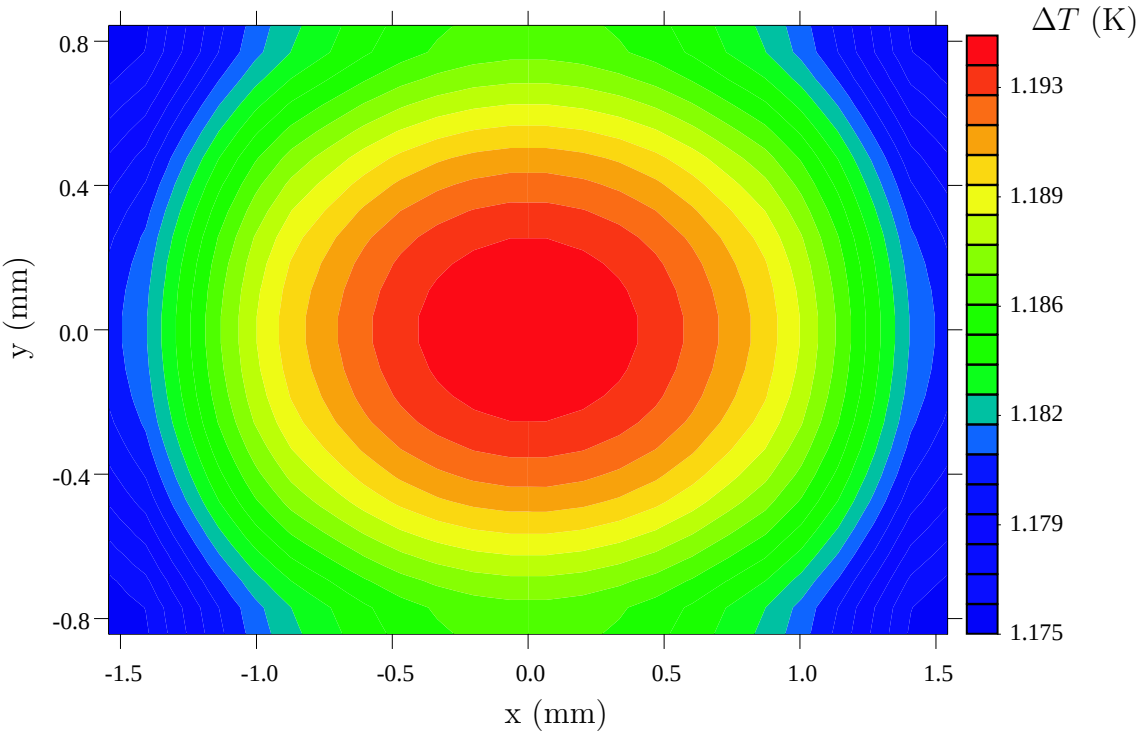

Figure 5.17: Steady-state temperature distribution in reference thermistor layer when instrument is viewing a blackbody source at 359 K

the temperature distributions are nearly identical to the respective temperature distributions in the reference detector.

The vertical temperature profile through the center of the active detector is shown in Figure 5.22 as a function of the distance from the detector surface. The maximum temperature is attained in the thermistor layer due to the electrical self-heating, and the temperature drops slightly towards the exposed surface due to the heat loss associated with radiative emission.

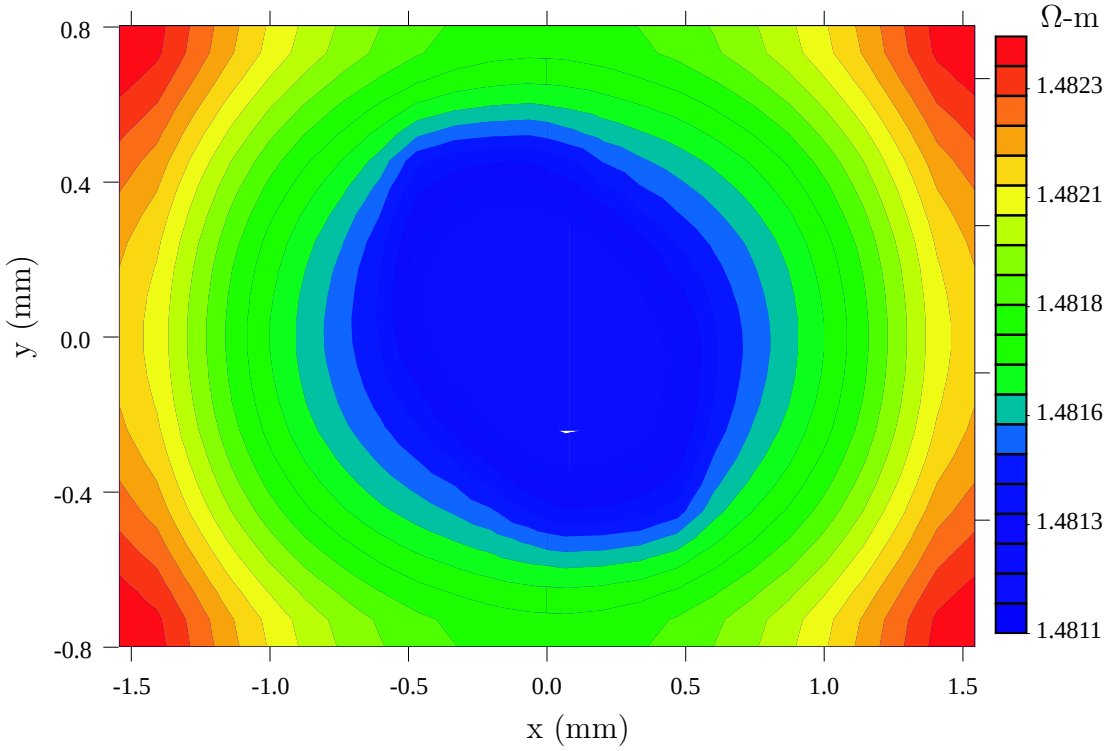

Figure 5.18: Steady-state spatial resistivity distribution in active thermistor layer when instrument is viewing a blackbody source at 359 K

5.4.4 Instrument output

The transient digital instrument response to a step input of 300 W/m^2· sr, corresponding to a blackbody source at 359 K, is shown in Figure 5.23. The steady-state instrument output is 1062 digital counts, which corresponds to a responsivity of 34 V/W. The time constant, τ, defined as the time necessary to achieve 63 percent of the steady-state response, is approximately 8 ms. The time constant compares well with measured CERES data and published values from previous models of the CERES instrument [Haeffelin et al., 1997]. The responsivity is considerably lower than reported values of approximately 64 V/W, and may

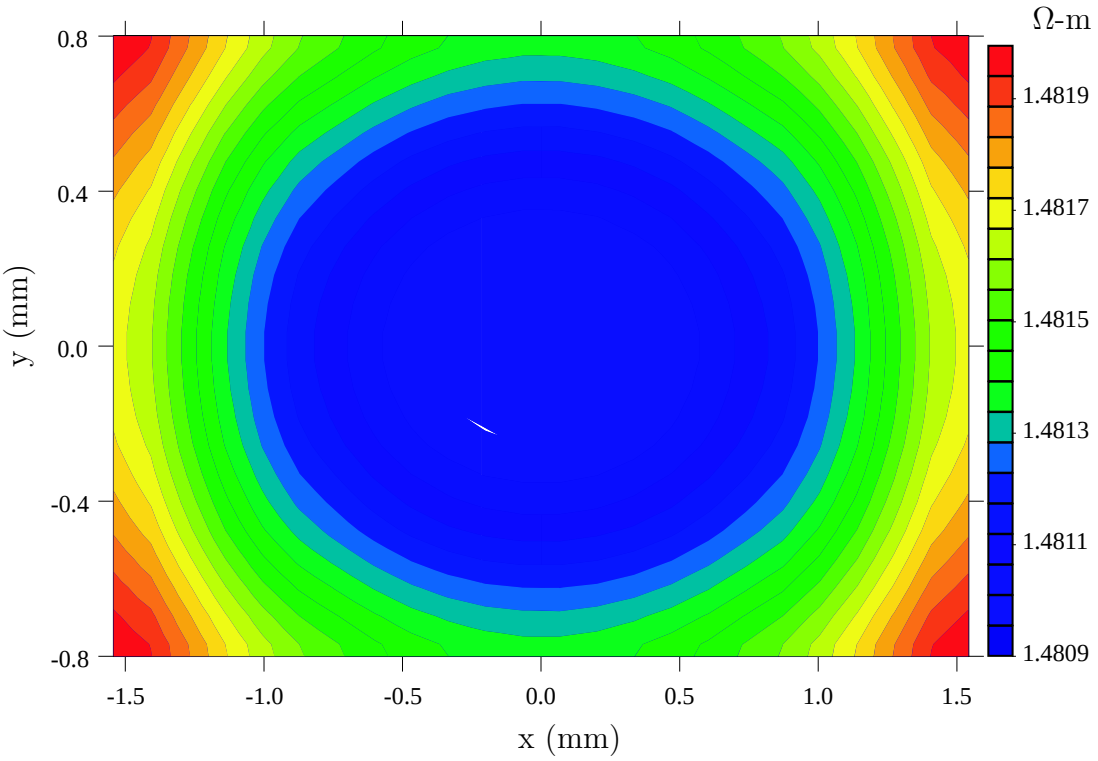

Figure 5.19: Steady-state spatial resistivity distribution in reference thermistor layer when instrument is viewing a blackbody source at 359 K

indicate that various model parameters need to be adjusted to more accurately match the instrument's performance.

5.5 Electrothermal sensitivity coefficients

To aid in detector design, it is important to understand the effect that the various different thermophysical properties have on the overall performance of the detector. It is possible to use sensitivity coefficients to quantify the sensitivity of the instrument output signal to various detector parameters. The sensitivity coefficient X_i is defined as the partial derivative

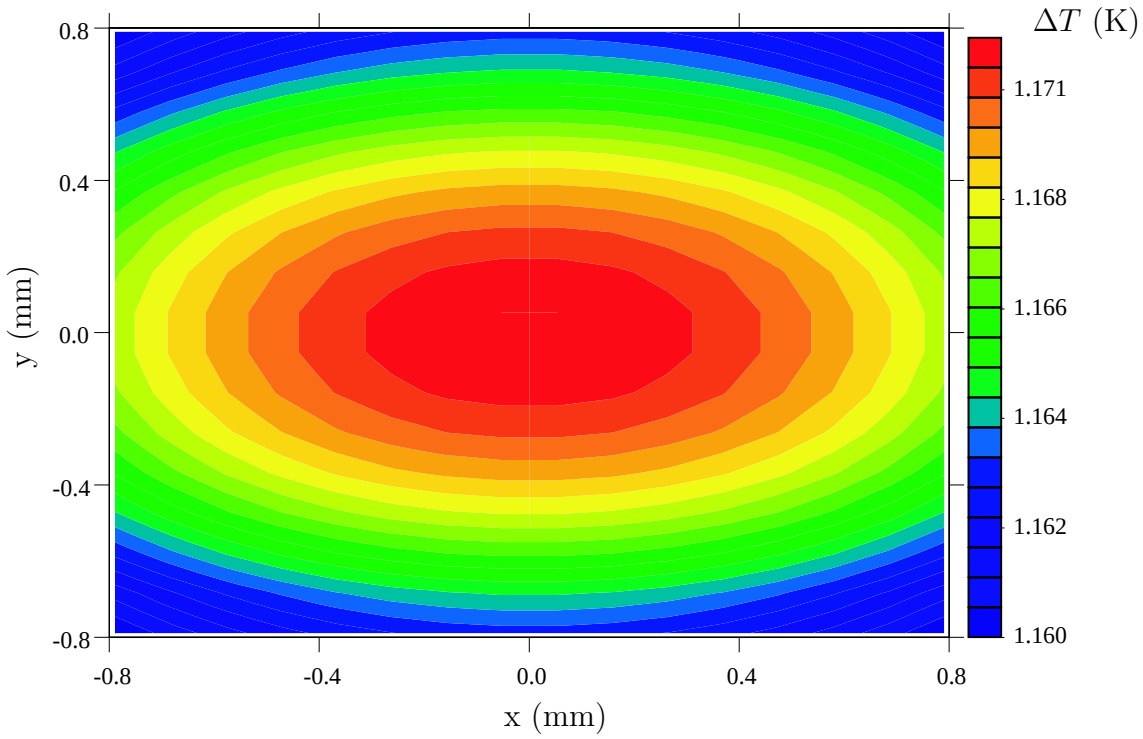

Figure 5.20: Steady-state temperature distribution on active flake when instrument is viewing cold space

of the digital instrument signal m with respect to the parameter of interest β_i,

$$X_i = \frac{\partial m}{\partial \beta_i} \tag{5.2}$$

It is convenient for purposes of intercomparison to normalize the sensitivity coefficients,

$$X_i^+ = \frac{\beta_{i,n}}{m_{ss}} \frac{\partial m}{\partial \beta_i} \tag{5.3}$$

where m_{ss} is the steady-state radiometric instrument output and $\beta_{i,n}$ is the nominal value of the parameter of interest. The normalized coefficients describing the sensitivity of the transient instrument output signal to the thermal conductivity of each material in the detector,

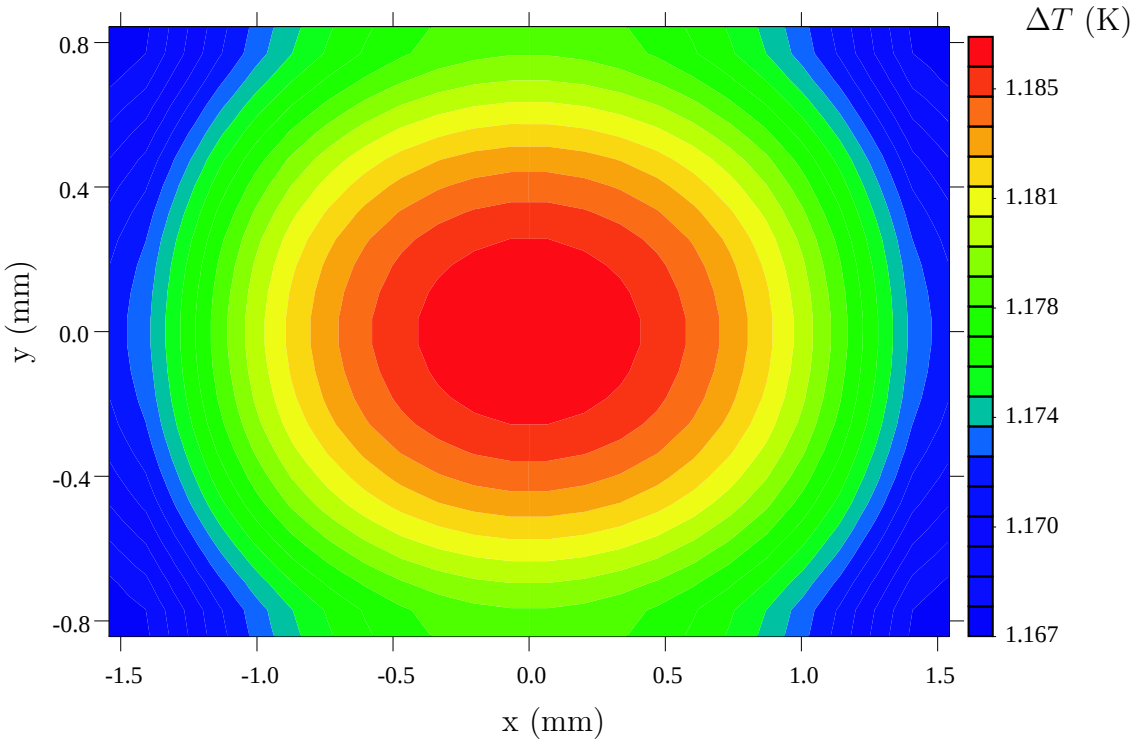

Figure 5.21: Steady-state temperature distribution in active thermistor layer instrument is when viewing cold space

$\beta_i = k_i$, are shown in Figure 5.24. The values of the sensitivity coefficients for the Kapton and epoxy layers below the thermsitor layer begin at zero at the initial time and increase to a steady-state value of approximately -0.1 for the epoxy layer and -0.8 for the Kapton layer. The layers above the thermistor layer begin at zero, peak at a small positive value at approximately 5 ms, and then decrease to a steady-state value of zero. Since the output signal is proportional to the temperature of the thermistor layer, an increase in the value of thermal conductivity in the layers below the thermistor have a large negative impact on the steady-state instrument output, with the Kapton having the largest impact. The thermal conductivity of the layers that are physically above the thermistor layer have some effect on the thermal time constant but no impact on the steady-state value, reflected in the fact that

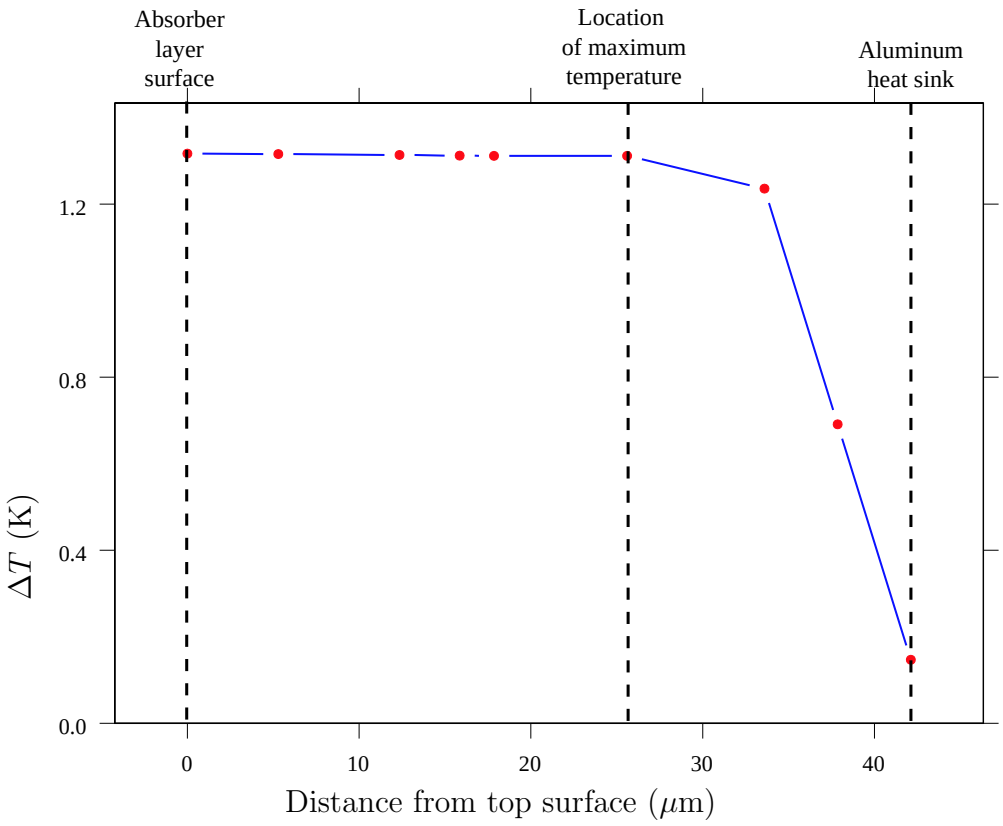

Figure 5.22: Steady-state vertical temperature distribution through center of the active detector when instrument is viewing cold space

the sensitivity coefficients reached a maximum value at 5 ms and then decreased towards zero as the system approaches steady-state.

The normalized coefficients describing the sensitivity of the transient instrument output signal to the thermal capacitance of each material in the detector, $\beta_i = \rho C_p$, are shown in Figure 5.25. Here the value of thermal capacitance in the thermistor layer has the largest impact on the output signal, with a maximum value of -0.225 at 8 ms and then decreasing to zero at steady-state. The other materials follow a similar profile, with smaller peak values ocurring at lesser time values. The layers above the thermistor layer all have a larger impact, and the Kapton and epoxy layers below the thermistor layers have little impact. The fact

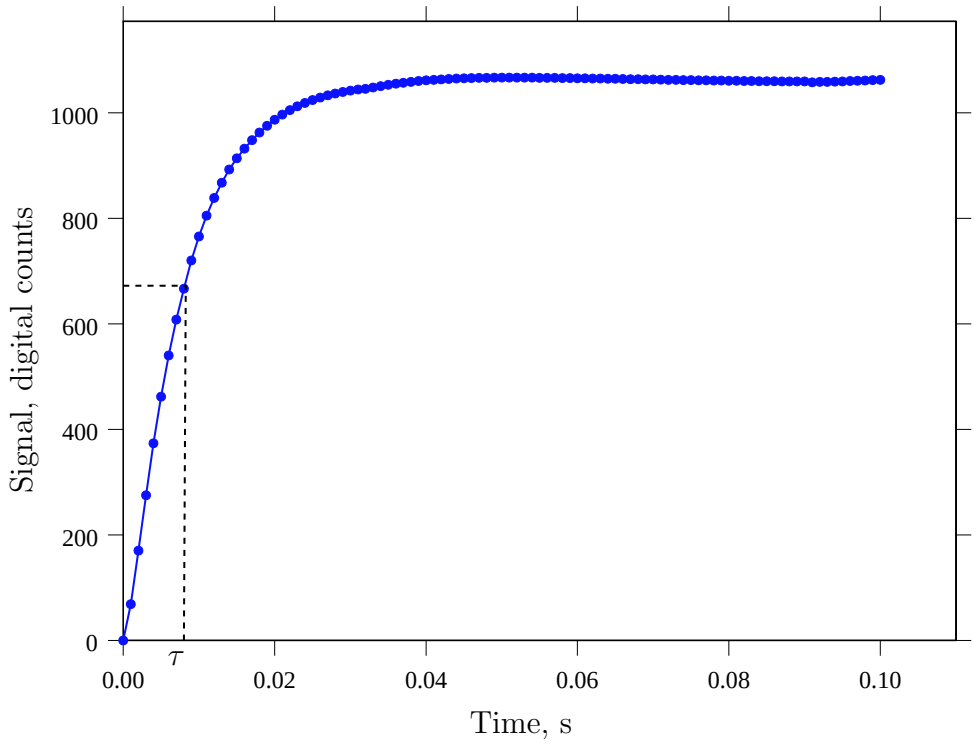

Figure 5.23: Evolution of instrument transient signal when viewing step-input blackbody source at 359 K

that all the sensitivity coefficients reach an early peak and tend to zero indicates that the thermal capacitance only affects the transient response of the instrument.

Another parameter that can affect the instrument output is the thickness, $\beta_i = \delta_i$, of each layer in the detector. Although not a thermophysical property, increasing the thickness effectively increases the thermal resistance and capacity and thus incorporates the effects of both the thermal conductivity and thermal capacitance. It is important because, from a design standpoint, the thickness of the layers is the easiest detector parameter to manipulate to achieve the desired instrument responsivity and time response. The normalized coefficients describing the sensitivity of the transient instrument output signal to the layer thickness of

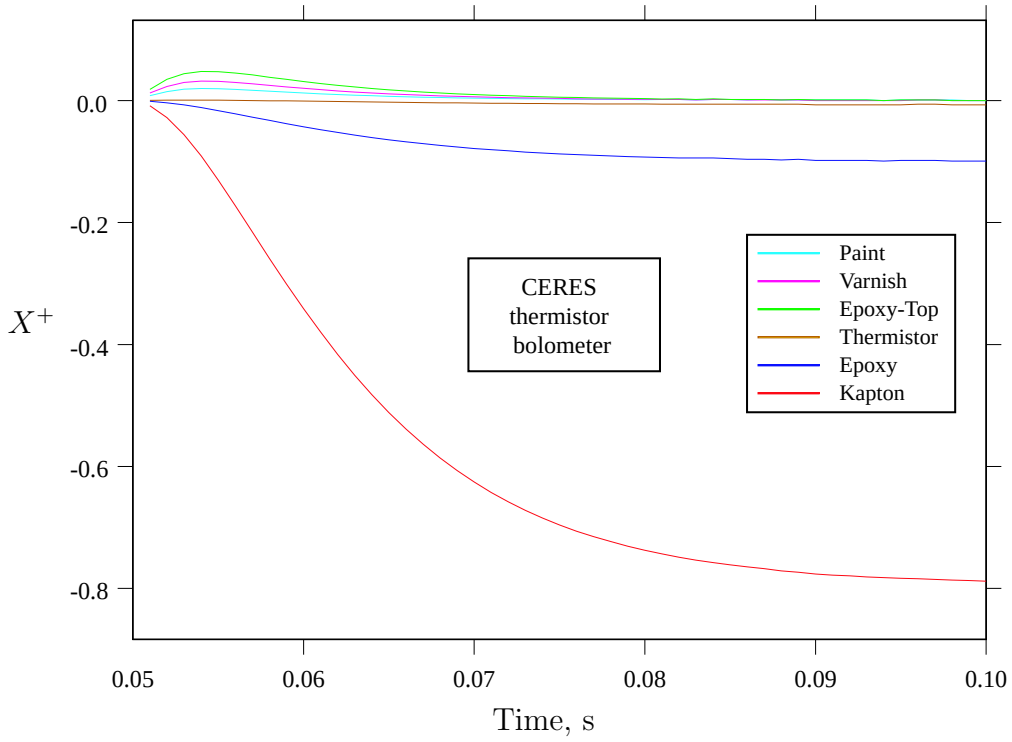

Figure 5.24: Sensitivity coefficients for thermal conductivity

each material in the detector are shown in Figure 5.26. The thickness of the Kapton layer has the largest impact on both the transient and steady-state response of the instrument. The thickness of the epoxy layers below the thermistor layer have a lesser inpact on the steady-state response and little impact on the transient response. The thickness of the thermistor layer has the largest impact on the transient response, followed in order of decreasing impact by that of the absorber paint and layer of epoxy above the thermistor.

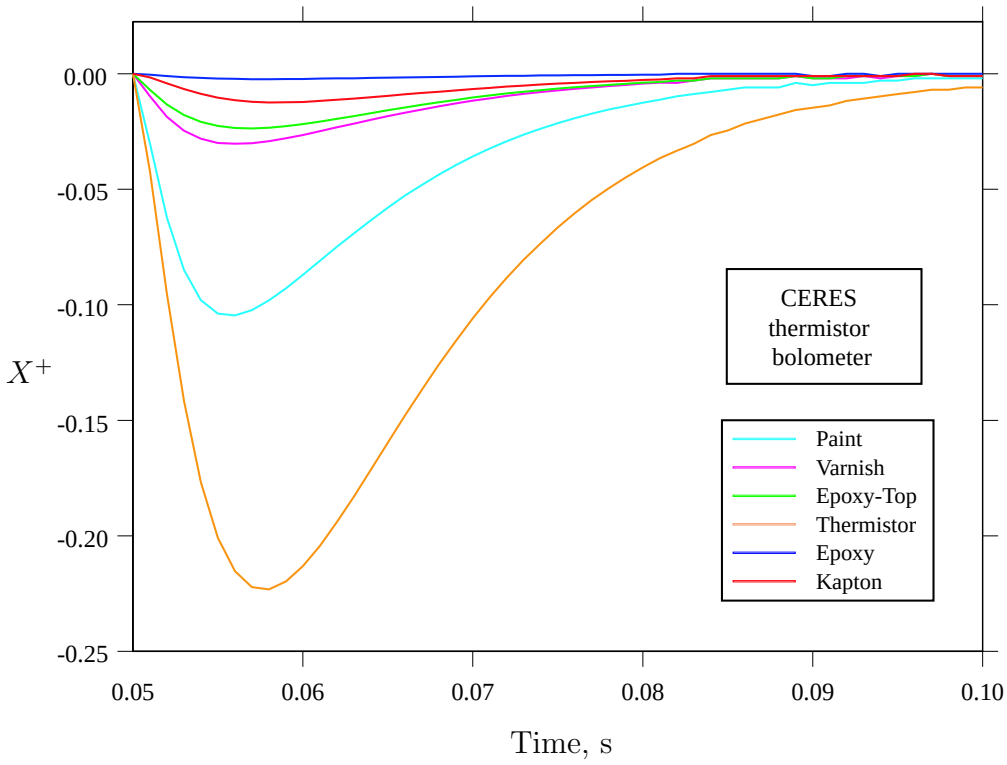

Figure 5.25: Sensitivity coefficients for thermal capacity

5.6 Matching electrothermal model to the actual CERES instrument

The value of the various optical and thermophysical properties and dimensions used in the end-to-end radiometric model are typically nominal values supplied by manufacturers, drawings, etc. Inherent uncertainties in the fabrication of the thermistor bolometer detectors, which is done by hand, can lead to actual parameter values that vary quite significantly from the nominal values. Also, contact resistance between the layers in the thermistor bolometer, both from fabrication and delamination effects, can cause increased thermal re-

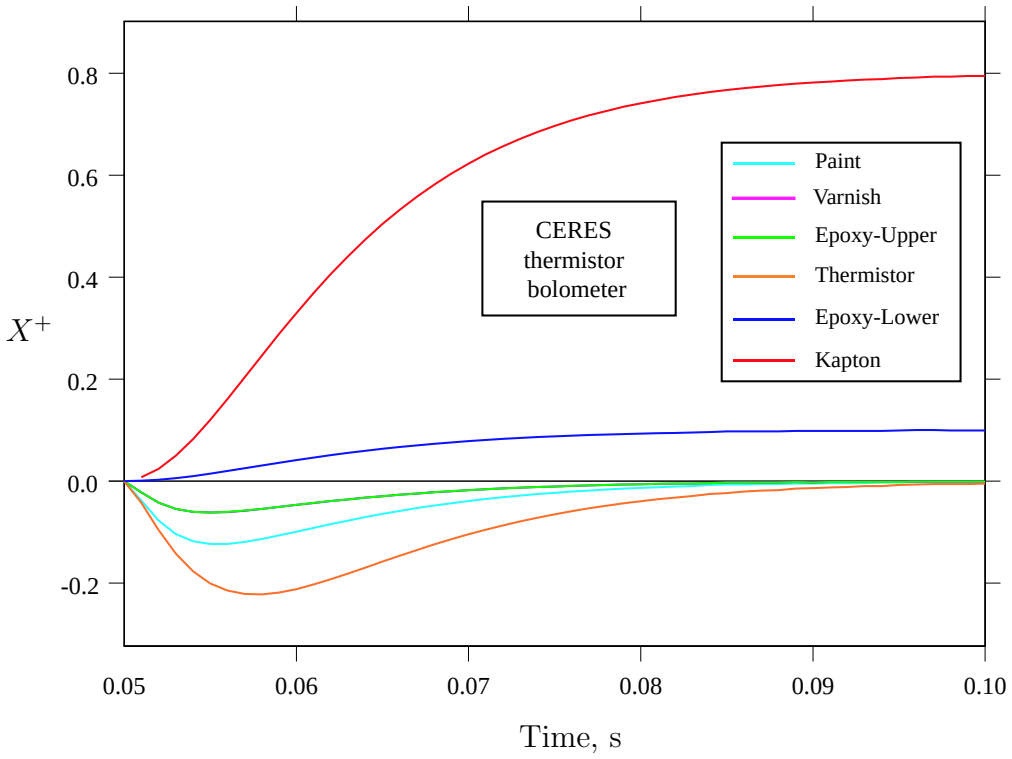

Figure 5.26: Sensitivity coefficients for layer thickness

sistance not accounted for in the nominal values. An interesting case in point is the reason the CERES detector assembly is constructed of two bolometers attached to separate aluminum substrates, which are then joined by an intervening layer of indium. In the initial attempts at constructing these detectors, two bolometers were laid down on opposite sides of a single substrate. Unfortunately, it proved impossible to achieve two detectors on a single substrate with matching responsivity and time response. It was then necessary to construct a large number of detectors on individual substrates, find two with closely matching values for responsivity and time response, and join the two substrates to form a single detector assembly. This illustrates how much variance from the nominal values exists in the actual instruments. While the nominal values are typically sufficient to obtain reasonable accuracy

from the model, optimal performance requires that the values of various model parameters be matched more closely with the values associated with the actual instrument, which are often unknown, at least to the desired precision.

One possible approach is to use parameter estimation techniques to estimate the values of the model parameters that results in the models performance most closely matching the observed behavior of the actual instrument. This can be done by varying the model parameters of interest in an intelligent manner that minimizes an objective function containing the measured output of the actual instrument and the output of the model [Beck and Arnold, 1977]. The varying of model parameters is usually done using some type of search or optimization algorithm, and the obective function to be minimized is the least squares error between the observed instrument output and the corresponding model output.

5.6.1 Description of genetic algorithms used in parameter estimation

The search algorithm used in the current application is genetic algorithms (GA's), which are probabilistic search algorithms based on evolutionary principles. The use of GA's in parameter estimation has been the subject of much study in recent years. García and Scott [1998a] performed a detailed study of the application of GA's to the simultaneous estimation of highly correlated thermal properties, presenting several case studies, and García [1999] used GA's to estimate the thermal and kinetic properties of polymers during curing. García and Scott [1998b] also used GA's to optimize experimental design for the estimation of thermal properties, and Hanuska [1998] used GA's to provide thermal characterization of complex aerospace structures comprised of several different materials.

The values of individual model parameters are represented as genes, and the array containing all the values of these genes is known as a chromosome, or indiviual. The initial population consists of a large number of individuals with values of their genes randomly generated within

the bounds of the search. Each individual is then passed to an objective function, where the values of each gene, which corresponds to a specific model parameter, are used to create an instrument model with those parameters. The model is then run and its output compared to the actual instrument output, and an error is calculated. This error is used to assign a fitness value to the individual, with a higher fitness value associated with a lower error. After all indiviuals have been assigned a fitness, certain individuals are selected to reproduce based on probablistic selection rules. Two individuals combine to produce offspring, with the values of the offspring's genes based on some combination of the values of the parents genes. A small percentage of the offspring are allowed to mutate, which means that one or more of their genes have their value slightly altered by some type of perturbation. The offspring are then inserted into the population, replacing some members of the previous population, and the cycle is started again. The rules governing the number of offspring and replacement of individuals by offspring vary among the different types of GA's. This process is allowed to iterate until some predetermined stopping point is reached, based on the number of generations or some convergence criteria. A more detailed explanation of the behavior of genetic algorithms is given in Chapter 3.

5.6.2 Model parameter estimation results

Three studies were performed to demonstrate the use of parameter estimation techniques to match the performance of a radiometric instrument model with an actual instrument, such as the CERES instrument. The studies used genetic algorithms to minimize the error between the instrument performance predicted by the model and the actual instrument performance. The instrument signal predicted by the model is compared with the measured instrument signal at each discrete time step, and the GA varies the instrument parameters in a probabilistic manner to minimize the cumulative error between the predicted and actual instrument response.

In each study, three GA runs were performed, with each run corresponding to a different set

of seeds used in the random number generator. A population of 100 chromosomes was used, and the GA ran for 800 generations using steady-state replacement at a rate of 10 individuals per generation. Binary tournament selection was used with random-point cross-over using a cross-over probability of 0.9; and creep-jump mutation was used with a creep factor of 0.05 and a creep probability of 0.5, and a mutation probability of 0.25. These settings were arrived at based on results of the study described in Chapter 6 and the author's experience using GA's. Each chromosome contained an array of real numbers containing values of the model parameters to be estimated.

The detector responsivity and time constant were used as measures of the detector performance for each chromosome. The three-dimensional transient electrothermal model of the detector assembly was used to provide values of the responsivity and time response. A coarse finite-difference discretization in the x-y direction was used to reduce computation time in the GA environment. This can be justified by the fact that the heat transfer in the detector is largely one-dimensional, with the two and three dimensional effects having little impact on the responsivity and time response of the detector, and by the fact that the GA's are known to provide very good results with approximate function evaluations Grefenstette and Fitzpatrick [1985]; Beasley et al. [1993a]. The exposed detector surface was represented by a 2×2 grid in the x-y direction; the epoxy, varnish, thermistor, and Kapton layers were represented by a 4×2 grid in the x-y direction; and the aluminum substrate was represented by a 6×4 grid in the x-y direction. The finite-difference discretization in the z-direction is the same as the end-to-end CERES instrument model.

The transient response of the CERES total channel, measured during instrument calibration, was used in the parameter estimation studies. The data, presented in Figure 5.27, was plotted as a normalized curve of the instrument's transient response. A responsivity of 62 V/W, determined from steady-state measurements, was also provided in the instrument calibration. By using a step input equivalent to a blackbody source at 359K as a reference, the given responsivity and the transient were used with a gain of 1760 from the electronics model to construct a transient signal representing the instrument's response. This signal was used as

the “measured” instrument response.

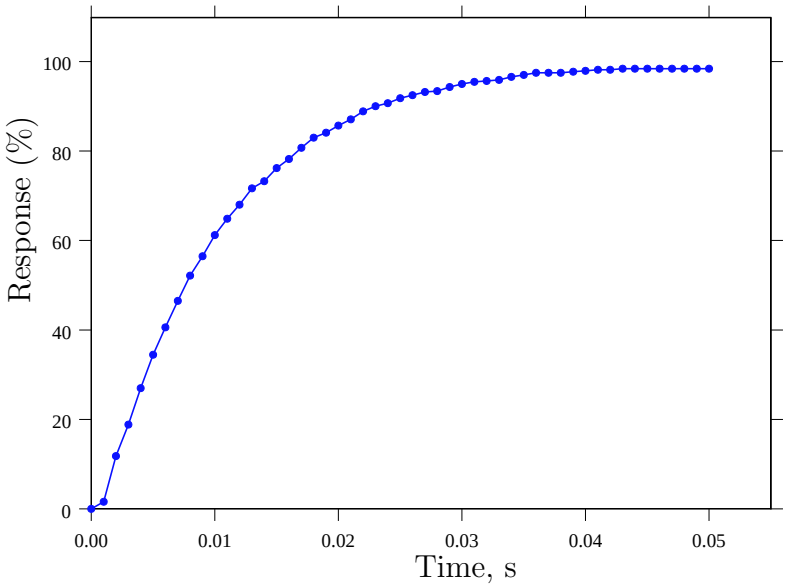

Figure 5.27: Measured transient response for CERES total channel

A step input heat flux boundary condition equivalent to the instrument viewing a 359 K blackbody source was applied, with 46.84 μW distributed over the detector surface according to the distribution factors, which were calculated beforehand and read from an input file for each model evaluation. The transient response of the instrument was calculated at each time step of 1 ms for a total time of 200 ms, with the step input applied at 100 ms, giving the detector system time to reach steady-state due to the electrical self-heating, thereby simulating the instrument’s space-look. The voltage difference from the bridge circuit at 100 ms was subtracted from the voltage difference at 200 ms, providing the voltage rise due to the step input. This voltage was divided by the total power absorbed on the detector flake (46.84 μW) and the resulting value interpreted as the detector responsivity. The objective function was comprised of a cumulative sum of the error between the measured CERES

output, (M_{meas}, and the output from the model,M_{model}, at each of N time steps,

$$E = \frac{1}{N}\sum(M_{meas} - M_{model})^2 \tag{5.4}$$

The fitness assigned to each chromosome was simply the inverse of the relative error, because a lower error is associated with a more fit chromosome.

The first study was performed to determine the layer thicknesses that, when used in the electrothermal detector model, provide a predicted instrument response that matches the actual measured instrument response. The thermistor, Kapton, and epoxy layers were selected, based on the results of the sensitivity analysis, as candidates to be varied. All other materials were held constant at their nominal values for layer thickness and thermal properties, listed in Table 5.6. The thermistor, Kapton, and epoxy layers had their thermal properties held constant at their nominal values and the thicknesses were varied in the GA.

Table 5.6: Nominal values for layer thickness, thermal conductivity, and thermal capacity

	z (μm)	k (W/m-K)	ρC_p (J/m^3-K)
Epoxy	1.0	0.13	1.55×10^6
Kapton	7.52	0.12	1.25×10^6
Thermistor	15.0	8.36	3.56×10^6
Paint	10.6	0.209	9.35×10^5
Varnish/epoxy	7.5	0.1	1.15×10^6
Gold	0.5	293	2.50×10^6
Aluminum	3668	237	2.44×10^6
Indium	2.54	80.86	2.76×10^6

The layer thickness values for the three GA runs are displyed in Table 5.7, with the resulting values for responsivity and time constant. All three runs were able to provide values for layer thicknesses that produced a predicted instrument response that closely matched the actual response. The thickness values for the epoxy and Kapton layers are both larger than their nominal values from Table 5.6, but both are quite reasonable. This could be due to contact

resistances that may exist between layers, which can be modeled by using thicker layers to account for the increased thermal resistance. The thickness value for the thermistor layer is slightly smaller than the nominal value of 15 μm, but is a reasonable value based on the difficulty in accurately fabricating the thermistor bolometers to the nominal specifications. A plot of the predicted instrument output for all three cases and the measured instrument response is given in Figure 5.28. The predicted signal for each of the three cases matches the measured signal reasonably well, but there is a fair amount of variation for the entire transient response. This could be because the upper and lower limits placed on the the layer thickness values prevented the GA from obtaining the ideal solution.

Table 5.7: Optimal layer thicknesses to match thermistor bolometer model with CERES bolometer performance

	Run 1	Run 2	Run 3
Epoxy (μm)	3.34	3.37	2.06
Kapton (μm)	11.80	11.39	14.04
Thermistor (μm)	11.08	10.25	10.62
Responsivity (V/W)	62.89	62.47	62.64
Time constant (ms)	12.16	11.61	11.81

The next study determined the values of thermal conductivity and thermal capacity that, when used in the electrothermal detector model, provide a predicted instrument response that matches the actual measured instrument response. The thermal conductivity and thermal capacity of the thermistor, epoxy, and Kapton layers were varied in the GA, while their thicknesses, along with the thickness and thermal property values of all other layers, were held constant at their nominal values. The resulting values for thermal conductivity and thermal capacity, along with the resulting values of responsivity and time constant, are shown in Table 5.8. The thermal conductivity values for epoxy range from 0.058 to 0.075, which is lower than the nominal value of 0.12. This can be explained again by the possibility of contact resistance, which can be modeled by an effective thermal conductivity and thermal capacity of the layer that account for the increased thermal resistance. The thermal conductivity of

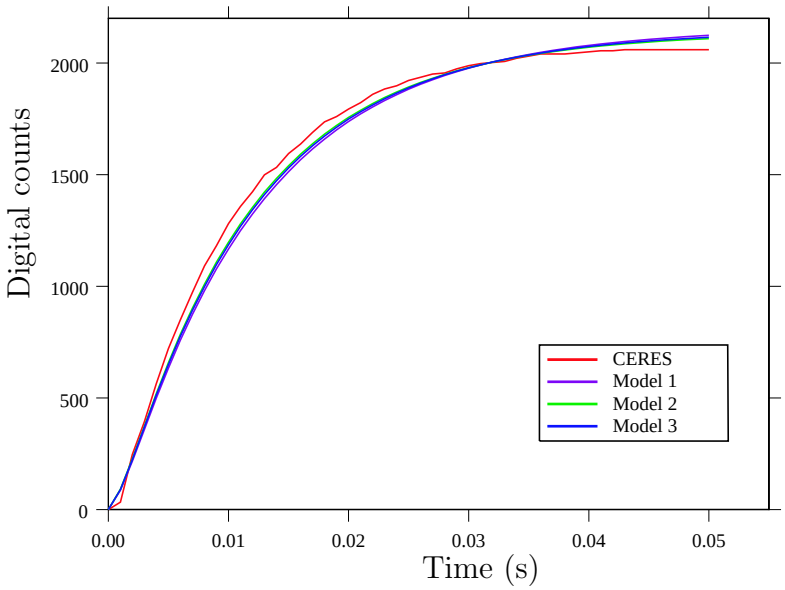

Figure 5.28: Measured transient response for CERES total channel and predicted model response for the case of varied layer thicknesses

Kapton ranged from 0.060 to 0.063, which is also lower than its nominal value, and can also be explained with the possibility of contact resistance. The thermal capacity of both epoxy and Kapton were much larger than their nominal values. While this can be explained in part by contact resistance, caution should be used when interpreting these results due to the very low sensitivity of the detector output to the thermal capacity of these two layers. The thermal conductivity of the thermistor layer ranged from 7.502 to 8.746, which is reasonably close to its nominal value, considering the low sensitivity coefficient associated with the thermsitor thermal conductivity. The thermal capacity of the thermistor layer ranged from 1.723 $\times 10^6$ to 2.081 $\times 10^6$, which is below the nominal value of 3.56 $\times 10^6$. This could not be explained by contact resistance, which would tend to increase the effective thermal capacitance of the layer. A plot of the predicted instrument output for all three cases and the measured instrument response is given in Figure 5.29. The predicted signal for each of

the three cases almost exactly matches the measured signal, with some slight variance at the first few and final few time steps.

Table 5.8: Thermal conductivity and thermal capacity values that match thermistor bolometer model with CERES bolometer performance

	Run 1	Run 2	Run 3
Epoxy: k (W/m-K)	0.058	0.071	0.075
Epoxy: ρC_p (J/m^3-K) $\times 10^6$	6.896	5.103	2.978
Kapton: k (W/m-K)	0.063	0.060	0.060
Kapton: ρC_p (J/m^3-K) $\times 10^6$	3.506	1.993	2.509
Thermistor: k (W/m-K)	7.796	7.502	8.746
Thermistor: ρC_p (J/m^3-K) $\times 10^6$	1.723	2.010	2.081
Responsivity (V/W)	61.53	61.45	61.46
Time constant (ms)	10.41	10.38	10.39

The final study investigated the combination of layer thickness and thermal property values that, when used in the electrothermal detector model, provide a predicted instrument response that matches the actual measured instrument response. The thermal conductivity, thermal capacity, and layer thickness of the thermistor, epoxy, and Kapton layers were varied in the GA, while the thickness and thermal property values of all other layers, were held constant at their nominal values. The resulting values for thermal conductivity and thermal capacity, along with the resulting values of responsivity and time constant, are shown in Table 5.9. The layer thickness of the epoxy and Kapton layers were close to their nominal values, varying between 2.127 and 3.032 μm for epoxy and 8.517 and 10.317 μm for Kapton. The layer thickness of the thermistor layer varied from 14.103 to 14.902 μm, which is is very close the nominal value of 15 μm. The thermal conductivity of epoxy varied between 0.084 and 0.104, and the thermal conductivity of Kapton varied from 0.091 to 0.104, both cases slightly below their respective nominal values. The thermal conductivity of the thermistor layer ranged from 6.110 to 8.129, which is slightly below the nominal value. This could again be attributed to the small sensitivity coefficient associated with the thermal conductivity. The thermal capacity of the epoxy and Kapton layers range from 3.504 $\times 10^6$ to 4.674 $\times 10^6$

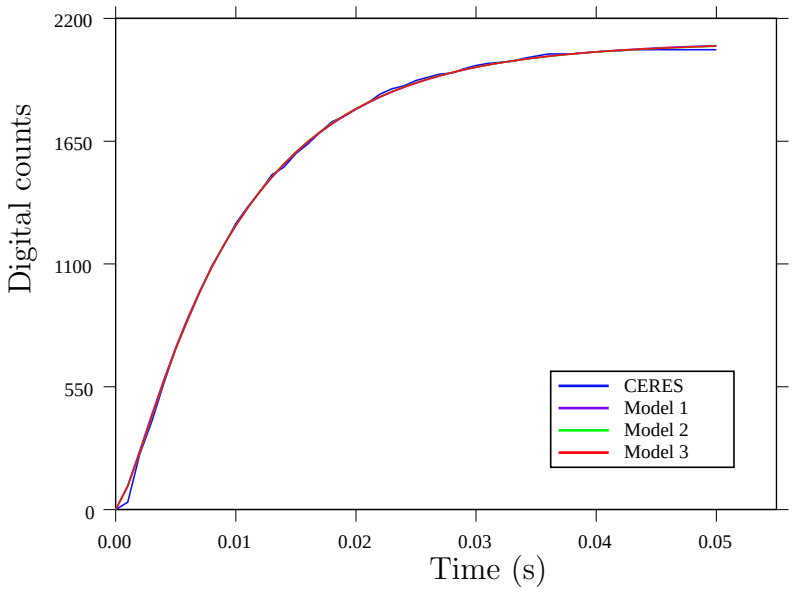

Figure 5.29: Measured transient response for CERES total channel and predicted model response for the case of varied thermal properties

and from 2.730 $\times 10^6$ to 4.223 $\times 10^6$, respectively. The values for both layers are well above their nominal values, which could be attributed to thermal contact resistance and small sensitivity coefficients. The value for thermal capacity of the thermistor layer varies from 1.487 $\times 10^6$ to 1.516 $\times 10^6$, which is again well below the nominal value. A plot of the predicted instrument output for all three cases and the measured instrument response is given in Figure 5.30. The predicted signal for each of the three cases almost exactly matches the measured signal, with some slight variance at the first few and final few time steps.

The results presented here indicate that the use of parameter estimation techniques in concert with a search algorithm such as GA's can help achieve optimal model performance. The optimal model parameters calculated in the three studies presented here varied somewhat from the nominal values given for the actual istrument, particularly in the case of the thickness and thermal capacity of the thermistor layer. Two possible explanations for this exist:

Table 5.9: Layer thickness, thermal conductivity and thermal capacity values that match thermistor bolometer model with CERES bolometer performance

	Run 1	Run 2	Run 3
Epoxy: z (μm)	3.032	2.127	2.824
Epoxy: k (W/m-K)	0.084	0.101	0.104
Epoxy: ρC_p (J/m^3-K) x 10^6	3.504	4.674	4.600
Kapton: z (μm)	8.517	10.317	9.279
Kapton: k (W/m-K)	0.104	0.091	0.095
Kapton: ρC_p (J/m^3-K) x 10^6	4.223	3.015	2.730
Thermistor: z (μm)	14.446	14.902	14.103
Thermistor: k (W/m-K)	8.129	7.355	6.110
Thermistor: ρC_p (J/m^3-K) x 10^6	1.487	1.504	1.516
Responsivity (V/W)	61.64	61.60	61.54
Time constant (ms)	10.42	10.45	10.42

either the nominal properties do not accurately reflect the actual as-built instrument, or a physical process in the instrument is not being accurately modeled. Additionally, care must be taken in interpreting the results, as the GA will tend to find a set a parameters that match the instrument performance, regardless of whether the parameters are actually feasible. It is necessary to bound the GA in such a way that resulting parameters are physically viable in the real instrument. Regardless, these techniques provide a powerful tool in optimizing the accuracy and performance of high-level numerical models. The use of GA's, in conjunction with high-level numerical models of the type presented here, presents a powerful tool for the design and optimization of radiometric instruments. This possibility is explored in the next chapter.

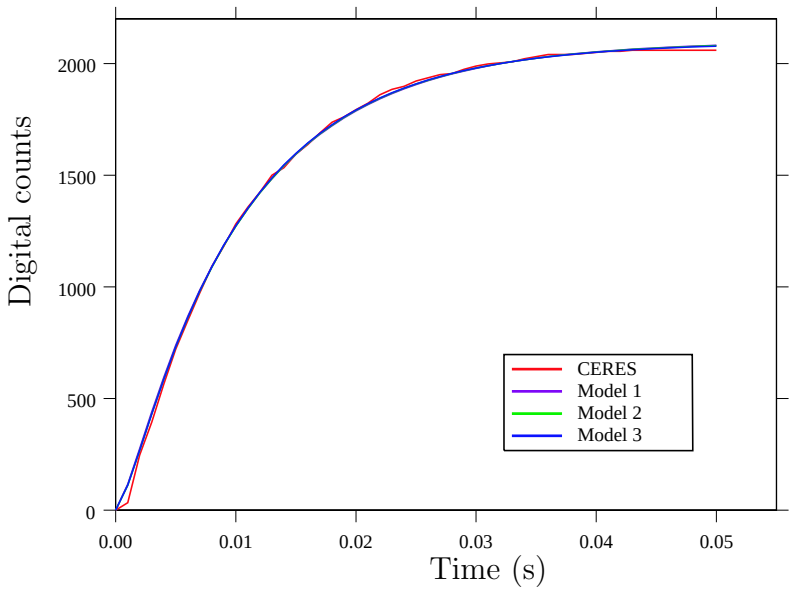

Figure 5.30: Measured transient response for CERES total channel and predicted model response for the case of varied layer thicknesses and thermal properties

Chapter 6

Use of genetic algorithms in radiometric instrument design and modeling

Genetic algorithms (GA's), presented in Chapter 3, are powerful probablistic search algorithms based on evolutionary principles. They can be applied to radiometric instrument design and modeling to provide a means for optimizing various performance parameters by determining the physical instrument parameters that result in optimal performance. Here they are applied to end-to-end radiometric instrument models to solve two types of problem: optimization of the optical system, and optimization of the detector assembly. For each case it is necessary to construct an objective function that represents the performance of the system, and then to determine which variables that objective function will be optimized against.

6.1 Optical systems

6.1.1 Objective function and optimization variables

The first step in using GA's to search for an optimal optical configuration is to define an appropriate objective function that describes the performance of the optical system. Two parameters can describe the performance of the optical system for incident collimated radiation: the size of the blur circle, and the optical throughput. These parameters can be considered separately or together in an aggregate objective function.

The size of the blur circle can be calculated by reading the x-y spatial coordinates of the energy bundles absorbed on the detector surface and expressing their number density distribution in terms of radial coordinates r and θ. The resulting values are then sorted by r in descending order, with the maximum values of r still containing energy bundles defining the radius of the blur circle. A flaw in this approach is the existence of stray energy bundles that can be absorbed well outside the radius of the blur circle, which could give the impression of a falsely large blur circle. This can be accounted for by rejecting a predetermined number of energy bundles with the largest values of r. By trial and error it was ascertained that, for the case of 10,000 rays emitted, rarely more than three of four stray energy bundles are absorbed. Accepting as the radius the radius of a circle that encloses 99 percent of the energy bundles is a reasonable approach to remove outliers and retain a good estimate of the size of the blur circle. The optical throughput can be estimated by the distribution factor between the entrance aperture and the detector surface, considering the detector as one surface element.

The goal of the optical system is to minimize the blur circle; therefore, an objective function representing the blur circle must return a value inversely proportional to the blur circle diameter. A simple approach is to return the inverse of blur circle diameter and assign that as the fitness for the associated chromosome in the GA algorithm. The optical throughput of the system, on the other hand, should be maximized, and an objective function representing only the optical throughput could simply return the value of the optical thoughput as the

fitness. Construction of an aggregate function is complicated by the fact that one parameter should be maximized and the other should be minimized. One approach is to return the root-square mean of the sum of the weighted inverse of the blur circle and the weighted optical throughput

$$Obj = \sqrt{w_1(B^*)^2 + w_2(O^*)^2} \tag{6.1}$$

where B^* is the normalized inverse of the blur circle diameter, O^* is the normalized optical throughput, and w_1 and w_2 are appropriate weight functions. Care must be taken in selecting the values to normalize the two parameters to ensure that they have approximately equal influence on the overall objective function, and the weight functions can then be used to vary the relative impact on the objective function. Using the CERES optical model with nominal parameters as a guide, 0.04 mm and 0.4 were taken as nominal values for the blur circle radius and optical throughput, respectively, and were used to normalize the two parameters in the aggregate objective function; i.e., $B^* = D_o/D$, where D_o=0.04 mm and $O^* = O/O_o$, where O_o=0.4.

Once an appropriate objective function has been created, the next step is to determine the model parameters to be used as search variables; i.e., the parameters encoded into the chromosomes of the GA. In the current application the location and physical properties of the active optical components (the mirrors and detector) are of primary interest, as they are the main components around which the optical system is designed. The model parameters that describe these components are the radius of the mirror slice a, the mirror depth c, the distance between the two mirrors, and the distance from the secondary mirror to the detector surface. By fixing the location of the detector and the mirror vertex of the primary mirror, both distances can be varied by changing the location of the secondary mirror. The search variables are then the radius of the mirror slice, the depth of both mirrors, and the location of the secondary mirror.

6.1.2 Approximate model study

Before the MCRT model of the optical system can be used in the GA environment, the number of emitted energy bundles to be used to evaluate each model must be determined. While an increasing number of energy bundles will provide more accurate estimates of the distribution factors, and thus of the optical performance of the telescope, it will also prohibitively increase the run-time of the GA. Fortunately, GA's have the useful property that they can obtain an optimal solution working with approximate function evaluations [Grefenstette and Fitzpatrick, 1985; Beasley et al., 1993a]. This feature admits the possibility of running the MCRT model with a relatively small number of energy bundles in the GA search while still obtaining an optimal solution. A study was performed to evaluate the variation in the objective function, using the aggregate function of Equation 6.1 with $w_1 = w_2 = 0.5$, to define the fitness of each chromosome. The results are presented in Table 6.1. The variation in the fitness with an increasing number of energy bundles is surprisingly small, with values of 7.805 when emitting a thousand energy bundles and 7.803 when emitting a million energy bundles. These results clearly indicate that emitting a small number of energy bundles, on the order of one thousand, will provide a sufficiently accurate estimate of the objective function to be used in the GA search.

Table 6.1: Sensitivity of the fitness of the optical model to the number of energy bundles emitted

Number of energy bundles	Fitness
500	7.80287
1000	7.80483
5000	7.80347
10000	7.80417
50000	7.80338
100000	7.80300
500000	7.80287
1000000	7.80284

6.1.3 GA parameter study

Prior to applying the GA to find the optimal optical system parameters, it is first necessary to fine-tune the GA to the problem at hand. A large number of parameters within the GA affect its performance, with a badly tuned GA performing as badly as a random search. A careful parametric study of the various settings of the GA is necessary to determine which parameter settings result in optimal performance.

A study of the various types of cross-over techniques was performed. Four cross-over mechanisms were considered: an arithmetic mean between the two parents, a fitness-weighted arithmetic mean, an arithmetic mean with a random component added or subtracted, and a random value between those of the two parents. These four mechanisms are described in more detail in Chapter 3. For each case, three GA runs were performed, with each run corresponding to a different set of seeds used in the random number generator. A population of 100 individuals was used, with steady-state replacement at the rate of 10 individuals replaced each generation, and the GA was run for 800 generations. Binary tournament selection was used with a selection probability of 0.9; a cross-over probability of 1.0 was used; and creep-jump mutation was used with a creep factor of 0.10, a creep probability of 0.8, and a mutation probability of 0.15. These parameter settings were selected based on the author's experience with GA's. The results of the study are presented in Table 6.2. The study shows that the worst performance (smallest fitness) is associated with the weighted arithmetic mean, with the arithmetic mean and the arithmetic mean with a random component performing at approximately the same level, and the random point approach outperforming (largest fitness) all other mechanisms. The probable explanation for the superior performance of the random point approach is that it can search the full range of searchspace between the two parents most effectively.

A study of the two most common selection techniques was also performed. Roulette wheel selection uses a probability proportional to the fitness for each individual to select individuals for mating, whereas binary tournament selects two individuals from the population, with all

Table 6.2: Study of crossover techniques

	Test 1 fitness		Test 2 fitness		Test 3 fitness		
Crossover method	Best	Avg.	Best	Avg.	Best	Avg.	Mean
Average	0.378	0.376	0.532	0.519	1.967	1.942	0.959
Weighted average	0.332	0.330	0.327	0.327	0.643	0.627	0.434
Random average	0.954	0.954	0.839	0.832	1.155	1.115	0.983
Random point	2.029	2.003	0.917	0.912	1.561	1.547	1.502

individuals having an equal probability of being selected, and then using the individual with the higher fitness of the two for mating. For each case, three GA runs were performed, with each run corresponding to a different set of seeds used in the random number generator. A population of 100 individuals was used, with steady-state replacement at the rate of 10 individuals replaced each generation, and the GA was run for 800 generations. The random-point cross-over technique was used with a cross-over probability of 0.9; and creep-jump mutation was used with a creep factor of 0.10, a creep probability of 0.8, and a mutation probability of 0.15. The results of the study are presented in Table 6.3. Over the course of the three runs for each case, roulette-wheel selection slightly outperformed tournament selection, although not by a significant margin. One might reasonably conclude that the two methods both work well on the problem, and that the choice as to which to use is arbitrary.

Table 6.3: Study of selection techniques

	Test 1 fitness		Test 2 fitness		Test 3 fitness		
Selection method	Best	Avg.	Best	Avg.	Best	Avg.	Mean
Roulette wheel	0.620	0.614	0.861	0.842	0.635	0.609	0.705
Binary tournament	0.550	0.541	0.869	0.856	0.555	0.547	0.658

A study was also performed of the number of individuals replaced per generation for steady-state GA's. The study was an attempt to determine, for a fixed number of function evalua-

tions, the balance of the number of generations the GA is allowed to run and the number of offspring produced in each generation. Too many offspring each generation results in wasted function evaluations because the same searchspace can be searched by multiple similar individuals, and too few offspring can result in undersampling the searchspace each generation. The study considered three cases in which 10, 25, and 50 offspring were replaced. For each case, three GA runs were performed, with each run corresponding to a different set of seeds used in the random number generator. A population of 100 individuals was used, with steady-state replacement at the rate specified for each case, and the GA was run for a number of generations equivalent to 8,000 function evaluations, i.e., 8000 evaluations of the MCRT model . The random-point cross-over technique was used with a cross-over probability of 0.9; and creep-jump mutation was used with a creep factor of 0.10, a creep probability of 0.8, and a mutation probability of 0.15. The results, presented in Table 6.4, seem to indicate that, of the three cases, replacing 25 individuals each generation results in the highest average final fitness over the three runs for a fixed value of 8,000 function evaluations.

Table 6.4: Study of replacement techniques

	Test 1 fitness		Test 2 fitness		Test 3 fitness			
Number replaced	Best	Avg.	Best	Avg.	Best	Avg.	NFE	Mean
10	0.606	0.600	0.840	0.824	0.548	0.548	8,000	0.665
25	0.667	0.640	0.331	0.325	1.796	1.693	8,000	0.886
50	0.588	0.498	0.323	0.323	0.606	0.606	8,000	0.476

A comparison was performed of two different mutation mechanisms. The study considered two types of mutation: creep-jump mutation, where the mutation is expressed as either a local pertubation or a jump to a random point in the searchspace, and Gaussian mutation, which is a local perturbation with a Gaussian distribution centered on the original value of the gene. For each case, three GA runs were performed, with each run corresponding to a different set of seeds used in the random number generator. A population of 100 individuals was used, steady-state replacement at the rate of 10 individuals replaced each generation, and

the GA was run for 800 generations. The random-point cross-over technique was used with a cross-over probability of 0.9, the case of creep-jump mutation used a creep factor of 0.10 and creep probability of 0.8. The case of Gaussian mutation used a value of 5 percent of the original value of the gene for the standard deviation. In both cases a mutation probability of 0.15 was used. The results of the study are presented in Table 6.5. The creep-jump mutation technique outperformed the Gaussian approach, most likely due to the ability to jump to a new point in the searchspace while still searching locally, while the Gaussian approach has a very limited method to jump to a new region and is therefore primarily restricted to local searches.

Table 6.5: Study of mutation types

	Test 1 fitness		Test 2 fitness		Test 3 fitness		
Mutation type	Best	Avg.	Best	Avg.	Best	Avg.	Mean
Creep-jump	0.587	0.569	0.366	0.366	0.735	0.716	0.563
Gaussian	0.469	0.438	0.428	0.421	0.532	0.525	0.476

Real-valued GA's have a tendency to perform better with higher rates of mutation than binary-encoded GA's. However, the question of identifying the optimal rate of mutation is still open, and tends to be problem specific. A study was performed to compare the performance of the GA with increasing rates of mutation. Mutation rates from 0.05 to 0.95 were considered. For each case, three GA runs were performed, with each run corresponding to a different set of seeds used in the random number generator. A population of 100 individuals was used, steady-state replacement at the rate specified for each case, and the GA was run for 800 generations. Roulette-wheel selection was used with random-point cross-over using a cross-over probability of 0.9; and creep-jump mutation was used with a creep factor of 0.10, a creep probability of 0.8, and the mutation probability specified in each case. The results are summarized in Table 6.6. The optimal mutation rate appears to be 0.50, which is higher than conventional wisdom would suggest; however, the results are far from definitive, as it is difficult to ascertain a clear trend from the results.

Table 6.6: Study of mutation rate

	Test 1 fitness		Test 2 fitness		Test 3 fitness		
Mutation rate	Best	Avg.	Best	Avg.	Best	Avg.	Mean
0.05	0.374	0.374	0.483	0.483	0.298	0.298	0.385
0.10	0.628	0.626	0.477	0.472	0.464	0.454	0.523
0.15	0.735	0.732	0.425	0.419	0.498	0.496	0.553
0.20	0.364	0.362	0.438	0.433	0.413	0.412	0.405
0.25	0.312	0.310	0.993	0.957	0.593	0.588	0.632
0.50	0.615	0.606	0.850	0.831	1.064	1.024	0.843
0.75	0.480	0.456	0.581	0.557	0.727	0.621	0.596
0.95	1.163	1.054	0.440	0.398	0.542	0.486	0.715

The results from a study of the different types of mutation indicate that creep-jump mutation is the best performer for the current application. Since the mutation is expressed as a combination of creep and jump mutations, with a probability associated with each, a study of these probabilities is necessary to determine the optimal setting. Three cases were considered, with creep probabilities of 0.3, 0.5, and 0.8. The probability for jump mutation is the creep probability subtracted from unity in each case. For each case, three GA runs were performed, with each run corresponding to a different set of seeds used in the random number generator. A population of 100 individuals was used, steady-state replacement at the rate specified for each case, and the GA was run for 800 generations. Roulette-wheel selection was used with random-point cross-over using a cross-over probability of 0.9; and creep-jump mutation was used with a creep factor of 0.10, a creep probability of specified in each case, and a mutation probability of 0.15. The results are summarized in Table 6.7. A clear trend can be seen from the results, as the highest performance is associated with a creep probability of 0.3, and decreasing performance as the creep probability increases. This seems to indicate that the jump mutation is more important in the current application than the creep, likely indicating that the local searchspace is effectively searched by the cross-over operator and the ability to jump to a new area in the searchspace is more critical.

Table 6.7: Study of creep probability

	Test 1 fitness		Test 2 fitness		Test 3 fitness		
Creep probability	Best	Avg.	Best	Avg.	Best	Avg.	Mean
0.3	1.362	1.347	0.645	0.637	0.705	0.703	0.904
0.5	0.585	0.578	0.551	0.537	0.519	0.519	0.552
0.8	0.364	0.362	0.438	0.433	0.413	0.412	0.405

Another aspect of creep-jump mutation is the range of the local creep mutation. In the current work this is expressed as a fraction of the original gene value, and is named the *creep factor*. A study was done to determine the optimal value of the creep factor, and a range of values from 0.005 to 0.2 was used. For each case, three GA runs were performed, with each run corresponding to a different set of seeds used in the random number generator. A population of 100 individuals was used, steady-state replacement at the rate specified for each case, and the GA was run for 800 generations. Roulette-wheel selection was used with random-point cross-over using a cross-over probability of 0.9; creep-jump mutation was used with a creep factor specified, a creep probability of 0.8, and a mutation probability of 0.15. The results, summarized in Table 6.8, indicate that a creep factor setting of 0.05 provides optimal performance for the GA.

Table 6.8: Study of creep factor

	Test 1 fitness		Test 2 fitness		Test 3 fitness		
Creep factor	Best	Avg.	Best	Avg.	Best	Avg.	Mean
0.005	0.592	0.592	0.232	0.229	0.481	0.481	0.435
0.01	0.513	0.513	0.265	0.265	0.310	0.309	0.363
0.05	1.373	1.373	0.321	0.318	0.438	0.438	0.711
0.1	0.834	0.834	0.530	0.530	0.446	0.444	0.603
0.2	0.513	0.513	0.409	0.407	0.404	0.404	0.442

6.1.4 Optimal mirror depths

Once the optimal parameter settings for the GA were determined, the GA was applied to the end-to-end radiometric model to aid in determining model parameter settings that correspond to the optimal instrument performance. The GA was applied to the optical and detector assembly models separately, with the objective function measuring the optical performance and detector performance, respectively.

The initial study performed was to determine, for a two-mirror system with fixed mirror vertices and mirror slice widths (Z_v and a, respectively, as defined in Section 4.1), the values of the mirror depths, c_1 and c_2, that result in optimal optical performance. In this example, the optical performance is defined in terms of the size of the blur circle on the detector surface and the optical throughput of the optical system.

The initial study considered a simplified optical system consisting of a forward baffle, a primary mirror, a secondary mirror, a precision aperture, and the detector surface, as shown in Figure 5.1. These components were isolated for study because they are the components in the optical path of the incident energy bundles and they are the surfaces having the primary impact on the optical performance. All other surfaces, such as the instrument housing, spider, etc., have only a secondary effect on the optical performance. The initial study can be viewed as a preliminary design, which can then be refined using the full optical model.

The GA was encoded with chromosomes containing two real values for the depths c_1 and c_2 of the primary and secondary mirrors, as shown in Figure 4.2, and a third value A quantifying the minor axis for the systems with elliptical mirrors. A Monte Carlo ray trace with 1000 energy bundles was used to provide approximate values for the blur circle diameter and the optical throughput. These values were used in the objective function to provide a fitness value for each chromosome. A study was done for each of the two-mirror systems described in Section 2.2, with two GA runs performed for each system. A population of 100 individuals was used with steady-state replacement at the rate of 10 replaced per generation, and the GA was run for 800 generations. Roulette-wheel selection was used with random-point cross-over

using a cross-over probability of 0.9; and creep-jump mutation was used with a creep factor of 0.05 and a creep probability of 0.3, and a mutation probability of 0.25.

Table 6.9: Mirror depth study results

	c_1 (mm)	c_2 (mm)	A (mm)	Blur Circle (mm)	Optical Throughput	Fitness
Cassegrain Test 1	1.941	2.452	-	0.002	0.741	12494.9
Cassegrain Test 2	1.936	2.347	-	0.002	0.739	16476.1
Modified Cassegrain Test 1	1.323	0.388	-	0.107	0.669	238.3
Modified Cassegrain Test 2	1.121	0.136	-	0.050	0.565	507.4
Gregorian Test 1	1.949	0.931	10.59	0.090	0.351	281.6
Gregorian Test 2	1.494	0.412	4.52	0.047	0.304	544.3
Dall-Kirkham Test 1	1.253	0.311	36.90	0.038	0.406	660.6
Dall-Kirkham Test 2	1.483	0.824	31.45	0.149	0.465	170.9
Ritchey-Chrétian Test 1	0.565	0.046	-	0.246	0.118	103.2
Ritchey-Chrétian Test 2	0.837	0.681	-	1.257	0.096	20.2

The blur circle diamter, optical throughput, and resulting fitness values based on Equation 6.1 with equal weighting between the blur circle and optical throughput for all cases are shown in Table 6.9. The resulting blur circle radius values attained were smaller than that of the CERES telescope value of 0.04 mm for the Cassegrain system, were at approximately the same level in at least one test for the modified Cassegrain, Gregorian, and Dall-Kirkham systems, and were much poorer for the Ritchey-Chrétian systems. The poor values in the Ritchey-Chrétian systems indicate the possibility that the search bounds were not properly set or the presence of local minima which produce values for the blur circle much larger than that of the global minimum.

The study was repeated, using the same optical parameters encoded on the chromosome, but using the full optical model to evaluate the optical system as opposed to the simplified model used in the previous study. The same settings were used for the GA, and the aggregate objective function, Equation 6.1, was used to quantify the optical performance, with equal

weighting between the blur circle and optical throughput. A Monte Carlo ray trace with 1000 energy bundles was used to evaluate the blur circle and optical throughput. The results are tabulated in Table 6.10. The resulting blur circle radius values using the full optical model were generally larger than those for the simplified model. This is due to the GA trying to minimize the blur circle while maximizing the optical throughput, and the resulting values represent the best compromise between the two outcomes. The optical throughput values compared well with the optical throughput of the CERES instrument optical model using nominal dimensions, which produces an optical throughput of 0.37. The blur circle of the Cassegrain and Gregorian systems are equal or superior to that of the CERES model, the modified Cassegrain had blur circles slightly inferior to that of the CERES model, and the blur circle diameters of the Dall-Kirkham and Ritchey-Chrétian systems were much larger than that of the CERES model.

Table 6.10: Mirror depth study results for full optical model

	c_1 (mm)	c_2 (mm)	A (mm)	Blur Circle (mm)	Optical Throughput	Fitness
Cassegrain Test 1	1.905	1.880	-	0.012	0.363	1.399
Cassegrain Test 2	1.880	1.651	-	0.021	0.31	0.618
Modified Cassegrain Test 1	1.372	0.483	-	0.142	0.3668	2.098
Modified Cassegrain Test 2	1.372	0.457	-	0.137	0.361	0.624
Gregorian Test 1	0.163	0.324	36.74	0.047	0.139	0.601
Gregorian Test 2	1.406	0.521	16.38	0.085	0.365	0.690
Dall-Kirkham Test 1	1.801	1.893	25.85	0.732	0.214	0.319
Dall-Kirkham Test 2	1.593	0.947	53.12	0.338	0.369	0.549
Ritchey-Chrétian Test 1	0.721	0.178	-	0.736	0.109	0.207
Ritchey-Chrétian Test 2	0.696	0.145	-	0.598	0.116	0.218

6.2 Detector assembly system

A study was performed to determine the suitability of applying GA's to optimize the performance of the thermistor-bolometer detectors. The thickness of the paint, thermistor, epoxy, and Kapton layers, of the detector were the detector parameters encoded into the GA chromosome. The detector responsivity and time constant were used as measures of the detector performance. The three-dimensional transient electrothermal model of the detector assembly was used to provide values of the responsivity and time response. A coarser finite-difference discretization in the x-y direction was used to reduce computation time in the GA environment. This can be justified by the fact that the heat transfer in the detector is largely one-dimensional, with the two and three dimensional effects having little impact on the responsivity and time response of the detector, and by the fact that the GA's are known to provide very good results with approximate function evaluations. The exposed detector surface was represented by a 2×2 grid in the x-y direction, the epoxy, varnish, thermistor, and Kapton layers were represented by a 4×2 grid in the x-y direction, and the aluminum substrate is represented by a 6×4 grid in the x-y direction. The finite-difference discretization in the z-direction is the same as the end-to-end CERES instrument model described in Chapter 5.

A step input heat flux boundary condition equivalent to the instrument viewing a 359-K blackbody source was applied, with 46.84 μW distributed over the detector surface according to the distribution factors, which were calculated beforehand and read from an input file for each model evaluation. The transient response of the instrument was calculated at each time step of 1 ms for a total time of 200 ms, with the step input applied at 100 ms, giving the detector system time to reach steady-state due to the electrical self-heating, and simulating the instrument's space-look. The voltage difference from the bridge circuit at 100 ms was subtracted from the voltage difference at 200 ms, providing the voltage rise due to the step input. This voltage was divided by the total power absorbed on the detector flake (46.84 μW) and the resulting value interpreted as the detector responsivity. The time constant used

is defined as the time it takes from the moment the step input is applied for the instrument output signal to rise to 63 percent of its steady-state value. Once the steady-state voltage of the detector model was calculated, the signal history was searched to find the signal matching 63 percent of the output signal, and the time associated with that signal was recorded as the time constant. Linear interpolation was used to calculate the time constant when the voltage fell between two recorded values.

The GA used consisted of chromosomes with four genes, each representing one of the four layer thicknesses under consideration. The objective function used was an aggregate function quantifying the effects of both the responsivity and time constant by

$$Obj = w_1 \frac{R}{\bar{R}} + w_2 \frac{\bar{\tau}}{\tau} \tag{6.2}$$

In Equation 6.2 $\bar{R}$ and $\bar{\tau}$ are reference values for responsivity and the time constant used to normalize the two quantities in the aggregate objective function, and w_1 and w_2 are weighting factors to vary the relative importance of responsivity and time constant in the objective function. Values of 50 V/W and 10 ms were used for the responsivity and time constant, respectively. Three studies were performed with weighting factors (w_1, w_2) of $(0.5, 0.5)$, $(0.75, 0.25)$, and $(0.25, 0.75)$. In each study, three cases were run using a population of 100 chromosomes for 800 generations using steady-state replacement at a rate of 10 individuals per generation. Roulette-wheel selection was used with random-point cross-over using a cross-over probability of 0.9; and creep-jump mutation was used with a creep factor of 0.05 and a creep probability of 0.5, and a mutation probability of 0.25. The maximum and minimum allowable values for the individual layer thicknesses are given in Table 6.11. The values for responsivity, time constant, and overall fitness for each case in the three studies are given in Table 6.12.

In the first study, with the weighting factors equal at 0.50 each, the results tended towards two very different sets of values, with one case converging on a set of layer thicknesses resulting in a responsivity of 80 V/W and time constant of 14 ms, and the other two cases converging on values resulting in a responsivity of approximately 26 V/W and time constant

Table 6.11: Maximum and minimum layer thickness values in GA search

	Paint (μm)	Thermistor (μm)	Epoxy (μm)	Kapton (μm)
Maximum	15.0	20.0	5.0	15.0
Minimum	5.0	10.0	0.1	5.0

Table 6.12: Study using genetic algorithms to find optimal layer thicknesses

w_1 (−)	w_2 (−)	GA Run	Paint (μm)	Thermistor (μm)	Epoxy (μm)	Kapton (μm)	R (V/W)	τ (ms)	Fitness (−)
0.50	0.50	1	7.80	10.51	4.82	14.04	80.0	14.3	1.149
0.50	0.50	2	5.87	10.24	0.49	5.40	26.1	4.9	1.275
0.50	0.50	3	6.73	10.89	0.34	5.62	25.8	5.2	1.224
0.25	0.75	1	7.89	10.42	0.68	6.33	30.4	5.9	1.415
0.25	0.75	2	6.79	10.78	0.62	5.26	26.4	5.3	1.553
0.25	0.75	3	7.81	12.70	0.83	5.70	29.1	6.4	1.315
0.75	0.25	1	7.33	10.53	4.74	14.44	80.7	14.3	1.385
0.75	0.25	2	10.36	10.73	4.58	14.71	80.4	15.3	1.369
0.75	0.25	3	8.86	11.13	4.63	14.42	79.4	15.0	1.358

of approximately 5 ms. This indicates the possibility of two local maximae in the searchspace that the GA population converges upon.

In the second study, the responsivity weighting factor and the time constant weighting factor were assigned values of 0.25 and 0.75, respectively. The population converged on layer thickness values that produced responsivity values ranging from 26 to 30 V/W and time constants ranging from 5.3 to 6.4 ms. The average thickness over the three cases for the paint layer was 7.5 μm, the average thermistor layer thickness was 11.3 μm, the average epoxy layer thickness was 0.7 μm, and the average Kapton layer thickness was 5.8 μm. These are all lower than the nominal thickness values for the CERES thermistor bolometer, which has a time constant of approximately 9 ms.

In the third study, the responsivity weighting factor and the time constant weighting fac-

tor were assigned values of 0.75 and 0.25, respectively. The population converged in layer thickness values resulting in responsivity values of approximately 80 V/W and time constant values ranging from 14.3 to 15.3 ms. The average thickness over the three cases for the paint layer was 8.9 μm, the average thermistor layer thickness was 10.8 μm, the average epoxy layer thickness was 4.7 μm, and the average Kapton layer thickness was 14.5 μm. The paint and thermistor layer thicknesses were not significantly different from the previous study; however, the thickness of the epoxy and Kapton layers is much greater than in the previous study. This is because the responsivity is most sensitive to changes in these two layers, and thus they saw the largest change when greater weighting was placed on the responsivity in the objective function.

The results presented in this chapter indicate that the use of GA's with high-level numerical models is a viable method to assist in the design of radiometric systems. One important result presented is that GA's can be used with MCRT models emitting only a small number of energy bundles, and still obtain optimal results. It had been believed that MCRT models were too computationally expensive to be useful in iterative search algorithms, which was a perceived limitation on their use that has been eliminated as a result of the present effort. It has been established that GA's can be used to find optimal optical configurations to optimize the performance of radiometric optical systems by minimizing the size of the blur circle on the detector while maximizing the optical throughput within the constraints of weighting factors placed on these two objectives. GA's were also used to vary the thickness and thermal properties of materials in the thermistor bolometer electrothermal model to achieve a high responsivity and small time constant.

The studies presented here are somewhat preliminary, in that they establish the viability of GA's as a radiometric instrument design tool. To achieve their full functionality, the GA's should be used with more complex radiometric system models, combining the optical and electrothermal models with models estimating weight, cost, time-to-market, and other performance parameters not discussed here. To achieve this, it seems likely that further effort will be required to reduce the computational time of the models for use in a probabilistic

environment such as GA's.

Chapter 7

Conclusions and Recommendations

7.1 Accomplishments and conclusions

The following accomplishments and conclusions can be drawn from the work presented in this dissertation:

- A three-dimensional optical design tool based on the Monte Carlo ray-trace method was developed to aid in the design and modeling of optical systems for use in radiometric instruments. The tool allows the user to build up a generic optical system by selecting one of five standard two-mirror configurations commonly used in space-based remote sensing applications, and to define the optical system by entering the spatial and optical characteristics that define the optical system.

- A three-dimensional transient electrothermal generic model of a thermistor bolometer was developed to aid in the design and modeling of thermistor bolometer thermal radiation detectors for radiometric applications. The model, based on the fully implicit finite difference method, allows the user to specify all spatial, thermal, and electrical properties, as well as the electronic components in the bridge circuit and amplifier system.

- The optical and electrothermal systems were integrated into a complete end-to-end radiometric instrument modeling environment. The model simulates the instruments dynamic optical and electrothermal response to produce the time-varying signal produced by the instrument based on specified spatially or temporally varying incident radiation.

- The end-to-end model was benchmarked with performance data and previous models of an operational CERES instrument total channel. A predicted responsivity of approximately 34 V/W and time constant of approximately 8 ms were achieved using nominal values. Transient data was combined with parameter estimation techniques to obtain layer thickness and thermal property values in the detector assembly that resulted in a predicted responsivity of 62 V/W and a time constant of 10.4 ms, which excactly matched the measured transient response. This establishes the validity of the modeling environment for use in simulating other instruments and as a radiometric design tool.

- Genetic algorithm (GA) search methods were combined with the optical and electrothermal radiometric systems models to develop a radimetric instrument design tool. The GA was used to determine the mirror slice depths for fixed mirror vertices and width that produced the optimal optical system performance for the five different two-mirror systems under consideration. In the limited studies presented, the GA provided equivalent or superior results for simplified models than the as-built CERES instrument in four out of the five cases. The results for the full models were not as satisfying, with only one case providing equivalent results to the CERES optical system consistently. This is likely due to the increased model complexity and limited computational times used.

- Low-level system models, including surprisingly undersampled MCRT models, can be used with GA's to establish optimal designs, because of the GA's ability to obtain optimal search results with approximate function evaluations.

- The use of GA's with MCRT and finite-difference models to achieve the optimal design of scanning radiometers has been demonstrated.

- The use GA's, in concert with parameter estimation techniques, presents a viable method to optimize model performance by determining the model parameters that most closely match the performance of the model with the performance of the system under study. This was demonstrated by determining the values for layer thickness, thermal conductivity, and thermal capacity of a numerical electrothermal model of a thermistor bolometer that matched nominal values of responsivity and time constant of an actual thermistor bolometer fabricated for use on the CERES instrument.

7.2 Recommendations for future work

Further areas of investigation extending the current work include:

- Extend the optical modeling environment to include one-mirror and off-axis optical systems

- Extend the detector modeling environment to include thermopile and pyroelectric detectors.

- Develop a Windows-based graphical user interface to facilitate the ease of use of the environment. While entering parameters via input files and text-based menus was sufficient for the scope of the research presented here, a GUI interface would greatly enhance the usability of the modeling environment.

- Pursue further work in incorporating genetic algorithms and other intelligent search methods in radiometric design. Incorporating more complex design variables such as accuracy of data produced, overall cost and weight, and spectral performance would be an interesting area for further work.

References

Antonisse, J., 1989, *A new interpretation of schema notation that overturns the binary encoding constraint*, in: **Proceedings of the Third International Conference on Genetic Algorithms**.

Astheimer, R. W., 1984, "Thermistor infrared detectors," Technical Memorandum, Barnes Engineering Company, Stamford, Connecticut.

Back, T., Fogel, D. B., Whitley, D., and Angeline, P. J., 2000, **Mutation operators**, Evolutionary Computation 1: Basic Algorithms and Operators, Institute of Physics Publishing, Bristol, UK.

Barkstrom, B., 1998, "Some thoughts on procedures for estimating measurement uncertainties in radiometric instruments," NASA Langley Research Center, Atmospheric Sciences Division.

Barkstrom, B. R. and Hall, J. B., 1982, *The Earth Radiation Budget Experiment (ERBE): an overview*, **Journal of Energy**, Vol. 6, pp. 141–146.

Barkstrom, B. R. and Smith, G. L., 1986, *The Earth Radiation Budget Experiment: Science Implementation*, **Review of Geophysics**, Vol. 24, pp. 379–390.

Barry, M., 1999, **Heat Transfer Issues in Thin-Film Thermal Radiation Detectors**, M.S. Thesis, Virginia Polytechnic Institute and State University.

Beasley, D., Bull, D. R., and Martin, R. R., 1993a, *An Overview of Genetic Algorithms: Part 1, Fundamentals*, **University Computing**, Vol. 15, No. 4.

Beasley, D., Bull, D. R., and Martin, R. R., 1993b, *An Overview of Genetic Algorithms: Part 1, Research Topics*, **University Computing**, Vol. 15, No. 2.

Beck, J. and Arnold, K., 1977, **Parameter Estimation in Engineering and Science**, John Wiley, New York, NY.

Blickle, T., 2000, **Tournament selection**, Evolutionary Computation 1: Basic Algorithms and Operators, Institute of Physics Publishing, Bristol, UK.

Bongiovi, R. P., 1993, **A parametric study of the radiative and optical characteristics of a scanning radiometer for Earth radiation budget applications using the Monte-Carlo method**, M.S. Thesis, Virginia Polytechnic Institute and State University.

Booker, L. B., Fogel, D. B., Whitley, D., Angeline, P. J., and Eiben, A., 2000, **Recombination**, Evolutionary Computation 1: Basic Algorithms and Operators, Institute of Physics Publishing, Bristol, UK.

Carman, S. L., Cooper, J. E., Miller, J., Harrison, E. F., and Barkstrom, B. R., 1992, *Clouds and Earth's Radiant Energy System (CERES)*, **Advanced Astrophysical Science**, Vol. 76, pp. 693–706.

Carnicero, B. A., 2001, **Characterization of Pyranometer Thermal Offset and Correction of Historical Data**, M.S. Thesis, Virginia Polytechnic Institute and State University.

CERES Science Team, 1995, "Clouds and Earth's Radiant Energy System (CERES) Algorithm Theoretical Basis Document," NASA Reference Publication 1376 Vol.2.

Chen, H., 1985, **Space remote sensing systems**, Academic Press, Inc., Orlando, Florida.

Chen, H., 1997, **Remote sensing calibration systems**, A. Deepack Publishing, Hampton, Virginia.

Coffey, K., 1998, **Next-Generation Earth Radiation Budget Instrument Concepts**, M.S. Thesis, Virginia Polytechnic Institute and State University.

Davis, L., 1991, **Handbook of Genetic Algorithms**, Van Nostrand Reinhold Publishing, New York, NY.

De Jong, K. A., 1992, *Genetic algorithms are not function optimizers*, **Foundations of Genetic Algorithms**, Vol. 2.

Dobarco-Otero, J., 2000, **Second-surface mirror effects in thin-film absorber layers**, M.S. Thesis, Virginia Polytechnic Institute and State University.

Doyle, J. F., 1995, *Determining the size and location of a crack in a frame structure*, in: **Proceedings of the Seventh Inverse Problems in Engineering Seminar, OSU, USA**.

Eshelman, L. J. and Schaffer, J. D., 1992, *Real-coded genetic algorithms and interval-schemata*, **Foundations of Genetic Algorithms**, Vol. 2.

Eskin, L., 1981, **Application of the Monte Carlo method to the transient thermal modeling of a diffuse-specular radiometer cavity**, M.S. Thesis, Virginia Polytechnic Institute and State University.

Fanney, A. H., 1975, **Experimental study of the LZEEBE measurement system under simulated flight conditions**, M.S. Thesis, Virginia Polytechnic Institute and State University.

Fonseca, C. M. and Fleming, P. J., 2000, **Multiobjective optimization**, Evolutionary Computation 2: Advanced Algorithms and Operators, Institute of Physics Publishing, Bristol, UK.

García, S., 1999, **Experimental Design Optimization and Thermophysical Parameter Estimation of Composite Materials Using Genetic Algorithms**, Ph.D. Thesis, Virginia Polytechnic Institute and State University.

García, S. and Scott, E. P., 1998b, *Use of Genetic Algorithms in Thermal Property Estimation: Part I - Experimental Design Optimization*, **Numerical Heat Transfer, Part A**, Vol. 33, pp. 135–147.

García, S. and Scott, E. P., 1998a, *Use of Genetic Algorithms in Thermal Property Estimation: Part II - Simultaneous Estimation of Thermal Properties*, **Numerical Heat Transfer, Part A**, Vol. 33, pp. 149–168.

Gen, M. and Cheng, R., 1997, **Genetic Algorithms and Engineering Design**, John Wiley and Sons, Inc., New York, NY.

Goldberg, D. E., 1989, **Genetic Algorithms in Search, Optimization, and Machine Learning**, Addison-Wesley.

Goldberg, D. E. and Deb, K., 1991, *A comparative analysis of selection schemes used in genetic algorithms*, **Foundations of Genetic Algorithms**, Vol. 1.

Grefenstette, J., 2000, **Proportional selection and sampling algorithms**, Evolutionary Computation 1: Basic Algorithms and Operators, Institute of Physics Publishing, Bristol, UK.

Grefenstette, J. J. and Fitzpatrick, J. M., 1985, *Genetic search with approximate function evaluations*, in: **Proceedings of the First International Conference on Genetic Algorithms**.

Haeffelin, M., 1993, **A numerical study of equivalence in scanning thermistor bolometer radiometers for Earth radiation budget applications**, M.S. Thesis, Virginia Polytechnic Institute and State University.

Haeffelin, M., 1997, **A Study of Earth Radiation Budget Radiometric Channel Performance and Data Interpretation Protocals**, Ph.D. Thesis, Virginia Polytechnic Institute and State University.

Haeffelin, M. P. A., Mahan, J. R., and Priestley, K. J., 1997, *Predicted dynamic electrothermal performance of thermistor bolometer radiometers for Earth radiation budget applications*, **Applied Optics**, Vol. 36, pp. 7129–7142.

Hajela, P. and Lin, C.-Y., 1992, *Genetic search strategies in multicriterion optimal design*, **Structural Optimization**, Vol. 4, No. 2.

Hancock, P. J., 2000, **A comparison of selection mechanisms**, Evolutionary Computation 1: Basic Algorithms and Operators, Institute of Physics Publishing, Bristol, UK.

Hanuska, A., 1998, **Thermal Characterization of Complex Aerospace Structures**, M.S. Thesis, Virginia Polytechnic Institute and State University.

Holland, J., 1975, **Adaption in Natural and Artificial Systems**, MIT Press.

Huang, B., 1990, *A Precise Measurement of Temperature Difference Using Thermopiles*, **Experimental Thermal and Fluid Science**, Vol. 3, No. 3, pp. 265–271.

Kandel, R., 1998, **L'Incertitude des Climats**, Hachette, Paris, France.

Korsch, D., 1991, **Reflective optics**, Academic Press, Inc., San Diego, California.

Kowsary, F., 1989, **Radiative Characteristics of spherical cavities having partially or completely specular walls**, M.S. Thesis, Virginia Polytechnic Institute and State University.

Lenoble, J., 1993, **Atmospheric Radiative Transfer**, A. DEEPACK Publishing, Hampton, Virginia.

Mahan, J., 2002, **Radiation Heat Transfer: A Modern Approach**, John Wiley & Sons, New York.

Mahan, J., Carnicero, B., Nevarez, F., and Sorensen, I., 1999, *Novel solar aureolemeter concept based on reflective optics*, **SPIE Proceedings, EOS/SPIE Symposium on Optical Systems Design, Berlin, Germany, May 25-28**, Vol. 3737.

Mahan, J. R. and Eskin, L. D., 1984, *The Radiation Distribution Factor - Its Calculation Using Monte Carlo and Example of Its Application*, **First UK National Heat Transfer Conference**, pp. 1001–1012, July 4-6, Leeds, Yorkshire, England.

Meekins, J. L., 1990, **Optical analysis of the ERBE scanning thermistor bolometer radiometer using the Monte Carlo method**, M.S. Thesis, Virginia Polytechnic Institute and State University.

Michalewicz, Z., 1996, **Genetic Algorithms + Data Structures = Evolution Programs**, Third Edition, Springer-Verlag, New York, NY.

Michalewicz, Z., 2000, **Hybrid methods**, Evolutionary Computation 1: Basic Algorithms and Operators, Institute of Physics Publishing, Bristol, UK.

Michalewicz, Z. and Janikow, C. Z., 1991, *Handling constraints in genetic algorithms*, in: **Proceedings of the Fourth International Conference on Genetic Algorithms**.

Michalewicz, Z., Nazhiyath, G., and Michalewicz, M., 1996, *A note on the usefulness of geometric crossover for numerical optimization problems*, in: **Proceedings of the Fifth International Conference on Genetic Algorithms**.

Miller, B. L. and Goldberg, D. E., 1995, "Genetic algorithms, selection schemes, and the varying effects of noise," IlliGAL Report No. 95009, Illinois Genetic Algorithms Laboratory, Department of General Engineering, University of Illinois at Urbana-Champagne.

Mouroulis, P. and MacDonald, J., 1997, **Geometrical optics and optical design**, Oxford University Press, New York, New York.

Nevárez, F., 2001, **MCRT Design**, Ph.D. Thesis, Virginia Polytechnic Institute and State University.

Passwaters, J. O., 1976, **Detailed thermal analysis of a thin-shell, spherical radiometer in Earth orbit**, M.S. Thesis, Virginia Polytechnic Institute and State University.

Pollock, D. D., 1985, **Thermoelectricity, Theory, Thermometry, Tool**, ASTM Special Technical Publication, American Society for Testing and Materials, Philadelphia, PA.

Priestley, K. J., 1997, **Use of First Principle Numerical Models to Enhance the Understanding of the Operational Analysis of Space-Based Earth Radiation Budget Instruments**, Ph.D. Thesis, Virginia Polytechnic Institute and State University.

Rasnic, R. L., 1975, **A thermal and kinematic study of a thin-wall spherical shell satellite**, M.S. Thesis, Virginia Polytechnic Institute and State University.

Richardson, J. T., Palmer, M. R., Liepins, G., and Hilliard, M., 1989, *Some guidelines for genetic algorithms with penalty functions*, in: **Proceedings of the Third International Conference on Genetic Algorithms**.

Sánchez, M. C., 1998, **Optical Analysis of a Linear-Array Thermal Radiation Detector for Geostationary Earth Radiation Budget Applications**, M.S. Thesis, Virginia Polytechnic Institute and State University.

Sánchez, M. C., 2002, **Uncertainty and confidence intervals in Monte Carlo ray-trace methods**, Ph.D. Thesis, Virginia Polytechnic Institute and State University.

Sánchez, M. C., Mahan, J., Ayala, E. A., and Priestley, K. J., 1999, *Tools for predicting uncertainty and confidence intervals in radiometric data products*, in: **SPIE Proceedings, International Symposium on Remote Sensing, Florence, Italy**, Vol. 3840.

Sánchez, M. C., Mahan, J., Nevárez, F. J., and Priestley, K. J., 2000, *Uncertainty and confidence intervals in optical design using the Monte Carlo ray-trace method*, in: **SPIE Proceedings, International Symposium on Remote Sensing, Barcelona, Spain**, Vol. 4169.

Sánchez, M. C., Mahan, J., Nevárez, F. J., and Sorensen, I. J., 2001, *Bounding uncertainty in Monte Carlo ray-trace methods*, in: **SPIE Proceedings, International Symposium on Remote Sensing, Toulouse, France**, Vol. 4540.

Sarma, J. and De Jong, K., 2000, **Generation gap methods**, Evolutionary Computation 1: Basic Algorithms and Operators, Institute of Physics Publishing, Bristol, UK.

Savransky, M., 1996, **A study of transient heat conduction and thermal noise in an Earth radiation budget radiometer**, M.S. Thesis, Virginia Polytechnic Institute and State University.

Schaffer, J. D., 1985, *Multiple objective optimization with vector evaluated genetic algorithms*, in: **Proceedings of the First International Conference on Genetic Algorithms**.

Schaffer, J. D., Caruana, R. A., Eshelman, L. J., and Das, R., 1989, *A study of control parameters affecting online performance of genetic algorithms for function optimization*, in: **Proceedings of the Third International Conference on Genetic Algorithms**.

Schwefel, H. P., 2000, **Advantages and disadvantages of evolutionary computation over other approaches**, Evolutionary Computation 1: Basic Algorithms and Operators, Institute of Physics Publishing, Bristol, UK.

Smith, A. M., 1999, **Prediction and measurement of thermal exchanges within pyranometer**, M.S. Thesis, Virginia Polytechnic Institute and State University.

Sorensen, I. J., 1998, **Analytical and Experimental Characterization of a Linear-Array Thermopile Scanning Radiometer for Geo-Synchronous Earth Radiation Budget Applications**, M.S. Thesis, Virginia Polytechnic Institute and State University.

Sparrow, E. and Cess, R., 1966, **Radiation Heat Tranfer**, John Wiley and Sons, New York.

Sparrow, E. M., Eckert, E. R. G., and Jonsson, V. K., 1962, *An enclosure theory for radiative exchange between specularly and diffusely reflecting surfaces*, **Journal of Heat Transfer**, Vol. 84, pp. 294–300.

Spears, W. M., 1992, *Crossover or mutation?*, **Foundations of Genetic Algorithms**, Vol. 2.

Syswerda, G., 1989, *Uniform crossover in genetic algorithms*, in: **Proceedings of the Third International Conference on Genetic Algorithms**.

Tira, N. E., 1987, **Dynamic simulation of solar calibration of the total, Earth-viewing channel of the Earth Radiation Budget Experiment (ERBE)**, M.S. Thesis, Virginia Polytechnic Institute and State University.

Tira, N. E., 1991, **A study of the thermal and optical characteristics of radiometric channels for Earth radiation budget applications**, Ph.D. Thesis, Virginia Polytechnic Institute and State University.

Villeneuve, P. V., 1996, **A Numerical Study of the Sensitivity of Cloudy-Scene Bi-directional Reflectivity Distribution Functions to Variations in Cloud Parameters**, Ph.D. Thesis, Virginia Polytechnic Institute and State University.

Walkup, M., 1996, **A Monte Carlo optical workbench for radiometric imaging system design**, M.S. Thesis, Virginia Polytechnic Institute and State University.

Weckmann, S., 1997, **Dynamic Electrothermal Model of a Sputtered Thermopile Thermal Radiation Detector for Earth Radiation Budget Applications**, M.S. Thesis, Virginia Polytechnic Institute and State University.

Whitley, D., 1989, *The GENITOR algorithm and selection pressure: why rank-based allocation of reproductive trials is best*, in: **Proceedings of the Third International Conference on Genetic Algorithms**.

Whitley, D., 1993, "A Genetic Algorithm Tutorial," Technical Report CS-93-103, Department of Computer Science, Colorado State University.

Wielicki, B., Barkstrom, B. R., Harrison, E. F., Lee, R. B., Smith, G. L., and Cooper, J. E., 1996, *Clouds and Earth's Radiant Energy System (CERES): an Earth Observing System Experiment*, **Bulletin of American Meteorological Society**, Vol. 77, pp. 853–868.

Wolfe, W. L., 1996, **Introduction to infrared system design**, Vol. TT 24 of *Tutorial texts in optical engineering*, SPIE-The International Society for Optics, Bellingham, Washington.

Wright, A. H., 1991, *Genetic algorithms for real parameter optimization*, **Foundations of Genetic Algorithms**, Vol. 1.

Wyatt, C. L., 1987, **Radiometric system design**, Macmillan Publishing Company, New York, New York.

Vita

Ira Joseph Sorensen was born May 26, 1972, in Reedley, California. He attended Selma High School, from which he graduated in 1990. He received a Bachelor of Science in Mechanical Engineering from California Polytechnic State University, San Luis Obispo, where he was a member of Tau Beta Pi and an officer of ASME, in 1996. He relocated to Blacksburg, Virginia, in 1997 to begin his graduate studies at Virginia Polytechnic Institute and State University. Over the course of his graduate studies, he served as a research assistant in the Thermal Radiation Group, as well as a NASA Langley Researcher Summer Scholar (LARSS), a NASA Graduate Student Researcher Program (GSRP) fellow, and a NASA Graduate Student Co-Op. He received his Master of Science in Mechanical Engineering in December, 1998, and obtained his Ph.D in December, 2002.

This dissertation was typeset with LaTeX 2_ε[1] by the author.

[1] LaTeX 2_ε is an extension of LaTeX. LaTeX is a collection of macros for TeX. TeX is a trademark of the American Mathematical Society.