12-2012

Development of a Novel Technique for Predicting Tumor Response in Adaptive Radiation Therapy

Rebecca Marie Seibert
rseiber1@utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Rebecca Marie Seibert entitled "Development of a Novel Technique for Predicting Tumor Response in Adaptive Radiation Therapy." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Nuclear Engineering.

J. Wesley Hines, Major Professor

We have read this dissertation and recommend its acceptance:

Chester Ramsey, Lawrence Townsend, Hairong Qi

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Development of a Novel Technique for Predicting Tumor Response in Adaptive Radiation Therapy

A Dissertation Presented for the
Doctor of Philosophy Degree

The University of Tennessee, Knoxville

Rebecca Marie Seibert
December 2012

ACKNOWLEDGEMENTS

I dedicate this dissertation to my mother, Sylvia Seibert. She was my teacher, personal shopper, cheerleader, psychiatrist, best friend, and most importantly, my hero. Her kindness, strength, and faith always inspired me. Her bravery while battling ALS taught me never to give up. I could have never completed this work without her. I miss her every day but I know she is in heaven watching over me.

I would also like to express my deep gratitude to the 2 gentlemen that were most influential in helping me finish this work: Dr. Wesley Hines and Dr. Chester Ramsey. Dr. Hines is my major advisor. His enthusiasm, guidance, understanding, and patience have been driving forces through my graduate career. His mentorship was paramount in providing a well-rounded experience consistent with my long-term career goals. He encouraged me to not only grow as an engineer but also as an instructor and an independent thinker. One simply could not wish for a better or more amicable supervisor. Dr. Ramsey is my mentor and also one of the most brilliant medical physicists I know. He is also an amazing teacher and friend. He has helped me tremendously in my career as well as in life. It is impossible to put down in words how much his support means to me, and while I cannot even begin to describe all of the amazing things he has done, suffice it to say that I will be forever indebted to him for his kindness. I can say with certainty that I hit the ultimate jackpot when it came to professors.

I would like to thank the other members of my dissertation committee, Dr. Lawrence Townsend and Dr. Hairong Qi, for their valuable input. I would also like to thank all other members of Dr.

Hines' research group and my fellow residents at Thompson Cancer Center for their contributions to this work.

I am also grateful to the entire faculty and staff, both past and present, at the nuclear engineering department of the University of Tennessee. Everyone I interacted with there showed such professionalism and kindness. The entire department fosters a spirit of learning and offered so much support to their students. It has been a pleasure to be part of such a wonderful department.

I would also like to thank my father and sister. They have always believed in me and uplifted my spirits if ever I was feeling low. Finally, I would like to thank my husband, Ryan Graciano, without whom my efforts would not have been so fruitful. He stood by me through all good and challenging times, and has always encouraged me in this endeavor (even if it was only so we could finally stop paying tuition).

ABSTRACT

This dissertation concentrates on the introduction of Predictive Adaptive Radiation Therapy (PART) as a potential method to improve cancer treatment. PART is a novel technique that utilizes volumetric image-guided radiation therapy treatment (IGRT) data to actively predict the tumor response to therapy and estimate clinical outcomes during the course of treatment. To implement PART, a patient database containing IGRT image data for 40 lesions obtained from patients who were imaged and treated with helical tomotherapy was constructed. The data was then modeled using locally weighted regression. This model predicts future tumor volumes and masses and the associated confidence intervals based on limited observations during the first two weeks of treatment. All predictions were made using only 8 days worth of observations from early in the treatment and were all bound by a 95% confidence interval. Since the predictions were accurate with quantified uncertainty, they could eventually be used to optimize and adapt treatment accordingly, hence the term PART (Predictive Adaptive Radiation Therapy).

A challenge in implementing PART in a clinical setting is the increased quality assurance that it will demand. To help ease this burden, a technique was developed to automatically evaluate helical tomotherapy treatments during delivery using exit detector data. This technique uses an auto-associative kernel regression (AAKR) model to detect errors in tomotherapy delivery. This modeling scheme is especially suited for the problem of monitoring the fluence values found in the exit detector data because it is able to learn the complex detector data relationships. Several AAKR models were tested using tomotherapy detector data from deliveries that had intentionally inserted errors and different attenuations from the sinograms that were used to develop the model. The model proved to be robust and could predict the correct "error-free" values for a

projection in which the opening time of a single MLC leaf had been decreased by 10%. The model also was able to determine machine output errors. The automation of this technique should significantly ease the QA burden that accompanies adaptive therapy, and will help to make the implementation of PART more feasible.

Table of Contents

LIST OF FIGURES

1. INTRODUCTION

Radiation therapy or radiotherapy is the use of high energy radiation to control or kill cancer cells and shrink tumors. It is the most commonly used method for treating cancer and is recommended for many types of cancer [Khan 2003]. Recent years have seen rapid advances in the equipment and delivery techniques associated with radiation therapy. These advances in technology have enabled the use of predictive and prognostic tools that have then been successfully employed in other industries. Specifically, it is now possible to acquire data during radiotherapy treatments that are well-suited for the nonparametric modeling scheme. Application of these nonparametric models has led to the creation of a new radiotherapy treatment paradigm, entitled PART (Predictive Adaptive Radiation Therapy).

1.1 BACKGROUND AND RESEARCH NEED

Radiation is a process in which energetic particles or waves travel through a medium or space [Hendee et al. 2005]. High dose radiation can be used to treat cancer and other illnesses. Radiation can be directly deposited into the patient's body via radioactive substances (brachytherapy) or externally via external beam radiotherapy (EBRT). In EBRT, the source of radiation is outside the patient (such as a radiation beam generated by a linear accelerator) and directed to the tumor site. The high-energy photons or electrons deposit the radiation to the region containing the tumor to destroy the cancer cells and, with careful treatment planning, ensure the surrounding normal tissues remain unaffected. In EBRT, no radioactive sources are placed inside the patient's body with EBRT.

Although EBRT has been in use for the last 100 years, there have been rapid technological advances in EBRT technology over the past 2 decades that have rendered it an even more viable treatment option. A notable improvement was the integration of three-dimensional (3D) anatomical information into the radiation therapy planning process. Another key advancement was the development of intensity-modulated radiation therapy (IMRT). In IMRT, the intensity of the beam across different fields of radiation being delivered is modulated so that each field may have any number of both low- and high-intensity radiation areas. This difference in intensity enables improved control of the dose distribution, resulting in even better preservation of normal tissue.

Another advancement that has helped shape the field of radiation therapy is image-guided radiation therapy (IGRT). Normal structures and tumors can be in entirely different positions from one treatment session to the next due to several reasons, including organ filling, tumor shrinkage, weight loss, or even subtle movements that occur while breathing. In IGRT, EBRT or more frequently IMRT is carried out with the aid of imaging equipment. Generally, CT scans or X-ray images are obtained in the treatment room just before or during administration of the radiation treatment. These scans are then compared with the reference imaging dataset that was acquired at the time of simulation. This comparison allows adjustments to be made to the patient's position and/or the radiation beams to help ensure the radiation is being delivered as the treatment plan intended.

On the whole, all of these advances have improved targeting of tumors by radiation, while simultaneously limiting the exposure of healthy tissues to radiation. It is noteworthy that none of

these advancements would have been possible without the parallel advances in computing technology. These sophisticated treatment techniques require extensive data processing, and in turn allow more data to be collected. The collected data provide valuable information regarding radiotherapy. With the correct analysis, the data collected from many IGRT systems can be used to determine the patient's response to treatment and the efficacy of the treatment device. The goal of this work was to find a way to model the data acquired during IGRT treatments to improve radiation therapy. The modeling of the IGRT data has led to the invention of predictive adaptive radiation therapy (PART).

Nonparametric models have gained popularity in the last decade due to significant advances in statistics, applied mathematics, and computer science. In nonparametric models, the model structure is not previously specified but is instead determined from data. The specific modeling schemes used in this dissertation will be discussed in greater detail in subsequent chapters. The models used in this work were employed with the goal of improving the reliability and efficacy of EBRT and aiding the implementation of PART.

1.2 ORIGINAL CONTRIBUTIONS

In this study, a new approach, PART, that has never been used before in radiation therapy is presented. Although adaptive radiation therapy has been in place for some time, the predictive component is original to this work. A summary of the original contributions to this dissertation is presented below.

- Development and assessment of a non-linear, nonparametric empirical model for actively predicting the tumor responses (and the associated confidence interval)

while the patient is undergoing radiation therapy. This model employs a database of past tumor responses to predict the response of new patients during treatment by using limited observations collected early in the treatment. Such prediction of tumor behavior can potentially be utilized to adjust the treatment course. Application of these modeling techniques is the foundation of PART.

- Development and assessment of a non-linear, nonparametric empirical method for automatically evaluating exit dosimetry on a helical tomotherapy system using auto-associative nonparametric modeling that is robust and has the capability to learn complex detector data relationships, even with detector data with a low temporal resolution and beam attenuation from the patient. This technique will aid in the implementation of PART by helping to automate the QA process.

1.3 DOCUMENT ORGANIZATION

This dissertation is organized into 5 major sections. This chapter is an introduction of the work and outlines the original contributions made by the author. Chapter 2 is a literature survey providing an overview of radiation therapy and predictive modeling. Next, the development and assessment of PART for predicting lung tumor response are presented in Chapter 3. Chapter 4 discusses auto-associative kernel regression and explains how it can model sinogram data of tomotherapy to serve as a prognostic tool and improve quality assurance. Chapter 5 concludes the dissertation and discusses recommendations for future work.

2. LITERATURE SURVEY

This section presents the results of a literature survey that was carried out to describe the need of nonparametric modeling in radiation therapy. The survey gives a more detailed overview of radiation therapy. It discusses some of the advances in radiation oncology that were mentioned in the previous section with greater focus on the techniques used in this research. Finally, the history of nonparametric models is presented with a discussion of how these models are currently used in the field of radiation therapy.

2.1 RADIATION THERAPY

The primary goal of radiation therapy is to permanently damage the DNA of cancer cells, thereby resulting in cell death. Radiation was applied in medicine soon after the discovery of X-rays by Roentgen in 1895 [Williams 1902]. However, a considerable limitation at the time was the absence of high-energy beams that could adequately treat deep-seated tumors. However, in the late 1950s to the early 1960s, Co-60 units and megavoltage linear accelerators began to be used to noninvasively irradiate non-superficial lesions [Suntharalingam et al. 2005]. Today, medical linear accelerators (commonly called linacs) use microwave technology to accelerate electrons in a wave guide. These electrons can be collimated and used to treat shallow lesions. In the case of deeper tumors, the electrons collide with a target metal having high atomic number (usually tungsten). The collisions elicit high-energy X-rays from the target, which then serves as the source of radiation. Karzmark's [1981] primer offers an in-depth description on the history and operating principles of linacs. Figure 1 shows the schematic of a present day linac *(Images courtesy of Elekta).*

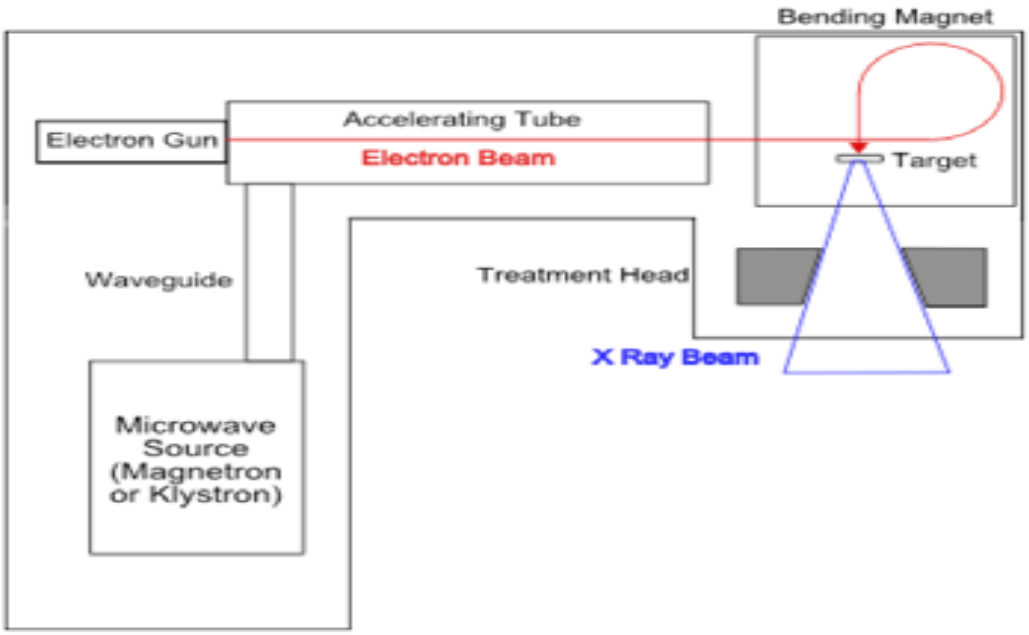


Figure 1 Diagram of a Linear Accelerator

For typical EBRT treatments, the patient is placed on a treatment couch while a mega-voltage X-ray beam is generated and emitted from the gantry head of the linac. The gantry can rotate isocentrically to treat the tumor from various angles. The prescription dose is usually divided into multiple fractions and delivered over a period of weeks rather than completed in a single undivided fraction. Dose fractionation is carried out in order to preserve normal tissue while allowing cell cycle redistribution and reoxygenation of the tumor cells for more effective tumor cell killing [Hall 2006]. The greatest challenge associated with radiotherapy is to minimize detrimental effects of radiation on the surrounding healthy tissue whilst simultaneously maximizing the degree of tumor destruction. The side effects from overdosing healthy tissue with radiation can be as severe as the effect of the tumor itself; conversely, reduced doses of irradiation may allow tumor regrowth. Many radiation delivery techniques have been proposed and used over the years, all with the aim to improve preservation of normal tissue and dose

conformity. The following sections describe some of the different techniques of radiation therapy, with emphasis on those pertinent to this research.

2.1.2 THREE-DIMENSIONAL CONFORMAL RADIATION THERAPY

Three-dimensional conformal radiation therapy (3D-CRT) is a sophisticated type of EBRT. In 3D-CRT, 3D images of the patient are acquired before the start of treatment. These images are used to delineate the tumor and surrounding normal structures. Generally, the physician contours and prescribes a dose to the planning target volume (PTV). The PTV is an expansion of the gross tumor volume (GTV), which is defined as the gross palpable or visible extent of the disease, and the clinical target volume (CTV), which contains the GTV plus an expansion to account for subclinical microscopic malignant disease [ICRU Report 50, 1978]. The PTV should account for all the possible geometrical variations and inaccuracies to ensure that the prescribed dose is actually absorbed in the CTV [ICRU Report 50, 1978]. From the 3D images, a computerized treatment planning system (TPS) renders a three-dimensional model of the treatment field and calculates how the dose will be deposited in the patient. The treatment planning system determines the shape of each treatment beam according to the projected target shape at a given gantry angle.

During delivery of the dose in 3D-CRT, the beam intensity of each treatment field is uniform, and only the beam aperture is shaped to match the projection of the target. The beam aperture can be modified using shielding blocks or a multi-leaf collimator (MLC) [Takahashi 1965]. The MLC is a device with movable leaves or shields of a material with a high atomic number; the device is attached to the gantry head [Brewster et al. 1995]. The leaves move independently to block some fraction of the radiation beam. Typical MLCs have 20–80 leaves, arranged in pairs. Computer controls are used to position many narrow, closely abutting leaves to generate a field

of arbitrary shape. Figure 2 illustrates this principle in a picture of an Elekta MLC device. Galvin et al. [1995] outlined the different design features of several MLC's and discuss how they are used as a substitute for alloy block field shaping. Since the early 1990s, 3D-CRT has been considered the standard of care in radiation therapy treatments for most tumors [Feng 2007]. However now that most linac designs employ MLCs, intensity modulated radiation therapy is becoming a more common delivery method.

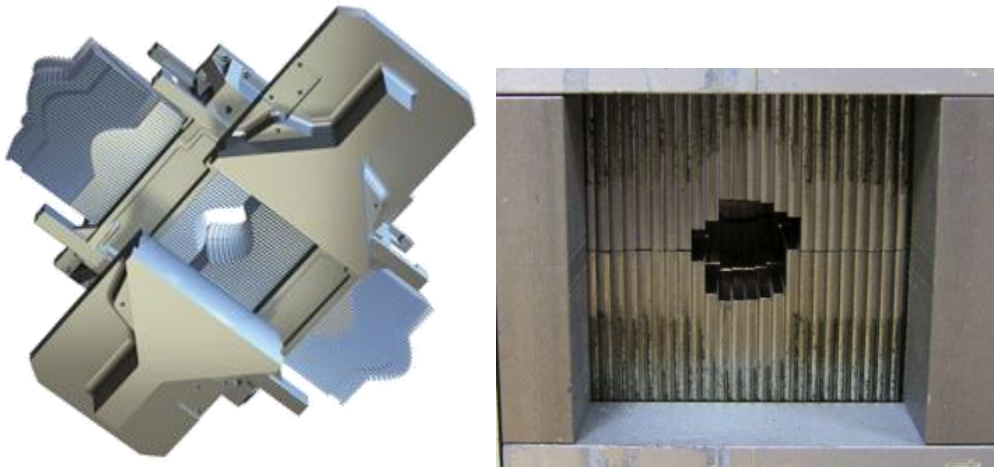


Figure 2 Elekta MLC and microleaf MLC

2.1.3 INTENSITY-MODULATED RADIATION THERAPY

Intensity-modulated radiation therapy (IMRT) is a relatively recent advance in radiation delivery. Unlike the 3D conformal technique, the beam intensity of IMRT varies across the treatment field depending on the shape of the desired target region. IMRT treatments aim to deliver an optimal dose to the target by modulating each incident treatment beam into smaller segments [Ezzell et al. 2003]. Rather than being treated with a large uniform beam, the patient is treated with many small pencil beams (referred to as beamlets), each of which can have a different intensity. The intensity of each beam is modulated by manipulating the MLC leaves moving into and out of the treatment field or with externally mounted 3-dimension compensators. By modulating the beam,

each beam delivers radiation with varying fluence. Thus, IMRT can deliver radiation with more degrees of freedom than 3D conformal therapy.

In IMRT, the treatment planning process involves selecting beam geometry and beamlet intensities to produce the best dose distribution. Since there are many possible beam geometries, beamlets, and range of beamlet intensities, there is an infinite number of possible treatment plans, and consistently and efficiently generating high-quality treatment plans is beyond human capability. Palta and Mackie [2003] explain how it is necessary to design and implement computer-based optimized decision support systems that can construct high-quality IMRT treatment plans in a short period of time. The increased complexity of IMRT planning and delivery gives rise to many safety concerns. The American Society for Radiation Oncology (ASTRO) Board of Directors recently commissioned a series of white papers addressing patient safety, with one specifically concentrating on IMRT [Moran et al. 2011]. This report, entitled "Safety considerations for IMRT: Executive summary," provides considerable background information and a detailed literature survey of IMRT. It also broadly addresses the safe delivery of IMRT, with a primary focus on recommendations for human error prevention and methods to reduce the occurrence of errors or machine malfunctions.

Safety is paramount in IMRT, especially because the segmentation of an individual treatment field leads to complex patterns of intensity distributions as well as non-trivial MU settings. Thus, it is recommended that patient specific QA be performed before the patient starts treatment. Generally, this QA process begins by using the treatment planning computer to recompute the individual delivery pattern onto a phantom. The treatment is delivered to the phantom and

analysis is done to verify that the delivered dose distributions agree to those that were produced by the treatment planning computer. Most phantoms designed for IMRT QA have a space to insert radiographic film (Figure 3). The film is calibrated so that its density corresponds to dose, allowing for comparison between planned and measured dose. Additionally some phantoms have a slot for an ion chamber so that absolute dose measurements can be taken. There are also surrogates for the complete set of measurements with film phantom/ion chamber. Examples of these surrogates include electronic portal imaging devices or electronic 2-D measuring tools, such as diode arrays or ion chamber arrays. AAPM Task Group 120 "Dosimetry tools and techniques for IMRT" provides a comprehensive overview of how dosimeters, phantoms, and dose distribution analysis techniques should be used to support the commissioning and quality assurance requirements of an IMRT [Ezzell et al. 2009]. Regardless of the device or method, it is necessary that every clinic that uses IMRT have a patient specific QA program which requires that the deliverability of the IMRT plan is validated before the patient undergoes treatment.

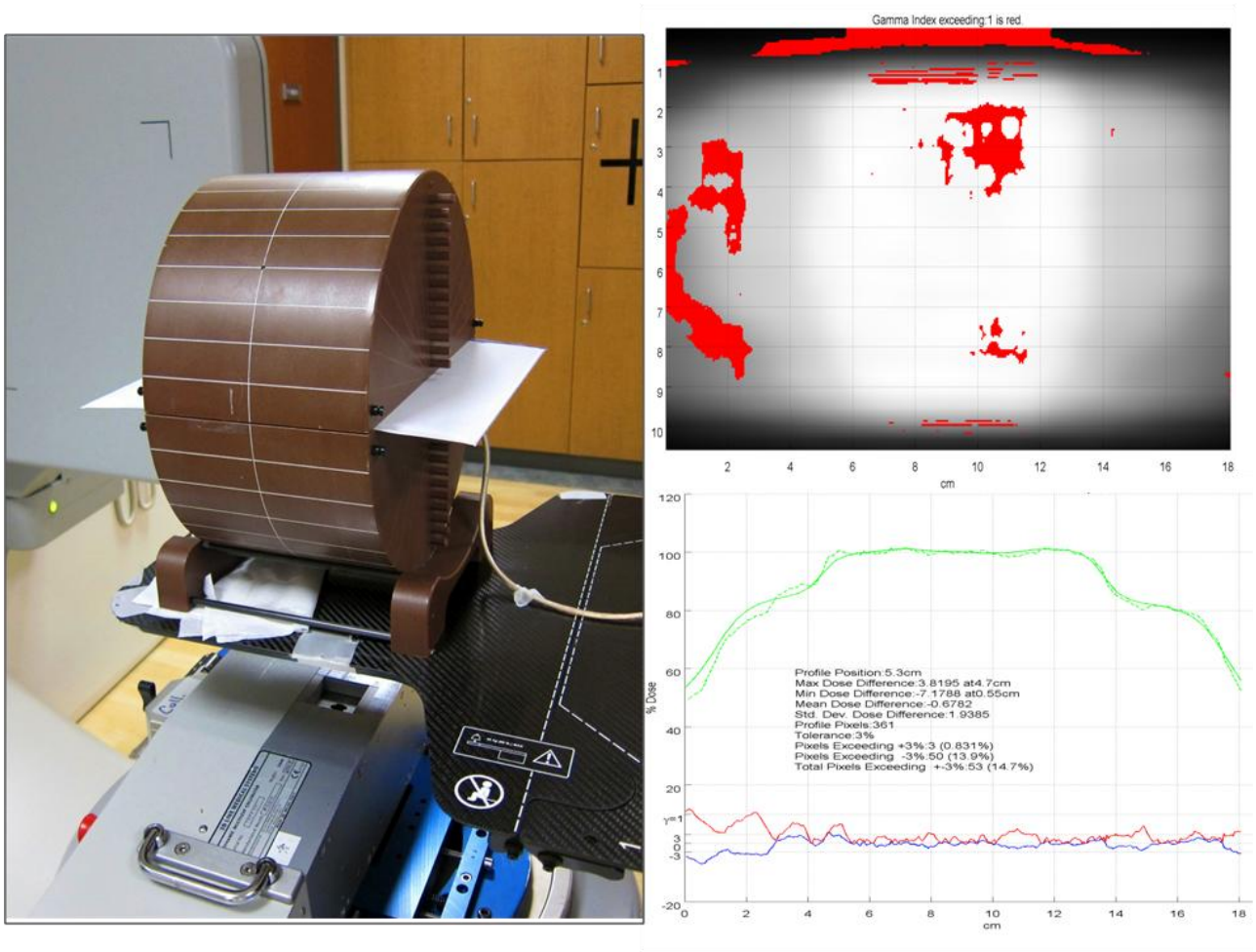


Figure 3 IMRT Film Phantom w/ Ion chamber (Left) and IMRT Film Analysis comparing Calculated vs. Measured Dose.

Although IMRT is much more time and resource intensive than conventional radiotherapy, it's many clinical benefits seem to be worth the extra effort. For many types of cancer, such as prostate cancer and head and neck cancer, the use of intensity modulation allows more concentrated treatment of the tumor volume, while limiting the radiation dose to adjacent healthy tissue. Veldeman et al. [2008] compared IMRT and non-IMRT treatments for different tumor sites. Since IMRT also facilitates directing higher radiation doses on regions within the tumor while sparing normal tissue, many physicians are treating tumors to a higher dose. However, this

dose escalation requires greater accuracy in patient positioning. The need for improved positioning accuracy encouraged the introduction of image-guided radiation therapy (IGRT).

2.1.4 IMAGE-GUIDED RADIATION THERAPY

IGRT is the use of frequent imaging during a course of radiation treatment to direct the radiation towards the target site. In IGRT, an image is acquired before treatment and compared to a reference image for dose planning. The results of this comparison are used to adjust the tumor position by isocenter correction. It is important to note the distinction of IGRT from the use of imaging to enhance target and organ delineation in the planning of radiation therapy. Basically, IGRT systems employ linear accelerators (or other radiation delivery devices) equipped with imaging technology that can image the tumor or internal anatomy in real-time before or even during the radiation therapy. Currently, various advanced imaging techniques are available for IGRT, such as transabdominal ultrasound [Langen et al. 2003, Fuss et al. 2004], implanted markers with kilovoltage (kV) or megavoltage (MV) X-ray images, kV computed tomography (kVCT) onrail [Wong et al. 2005], kV or MV cone beam computed tomography (CBCT) [Jaffray et al. 2002], and helical mega voltage CT (MVCT) [Kupelian et al. 2006].

Figure 4 illustrates a type of IGRT with a kV CBCT of a lung patient obtained using the patient in the treatment position directly before the beam was turned on. The CBCT is overlaid onto the reference simulation CT dataset. A checkerboard display is used where the darker image is the CBCT and the lighter image is the reference CT. This overlay helps clinicians to visualize whether or not the treatment volume is off target from the planned target and make the necessary adjustments to ensure accurate treatment.

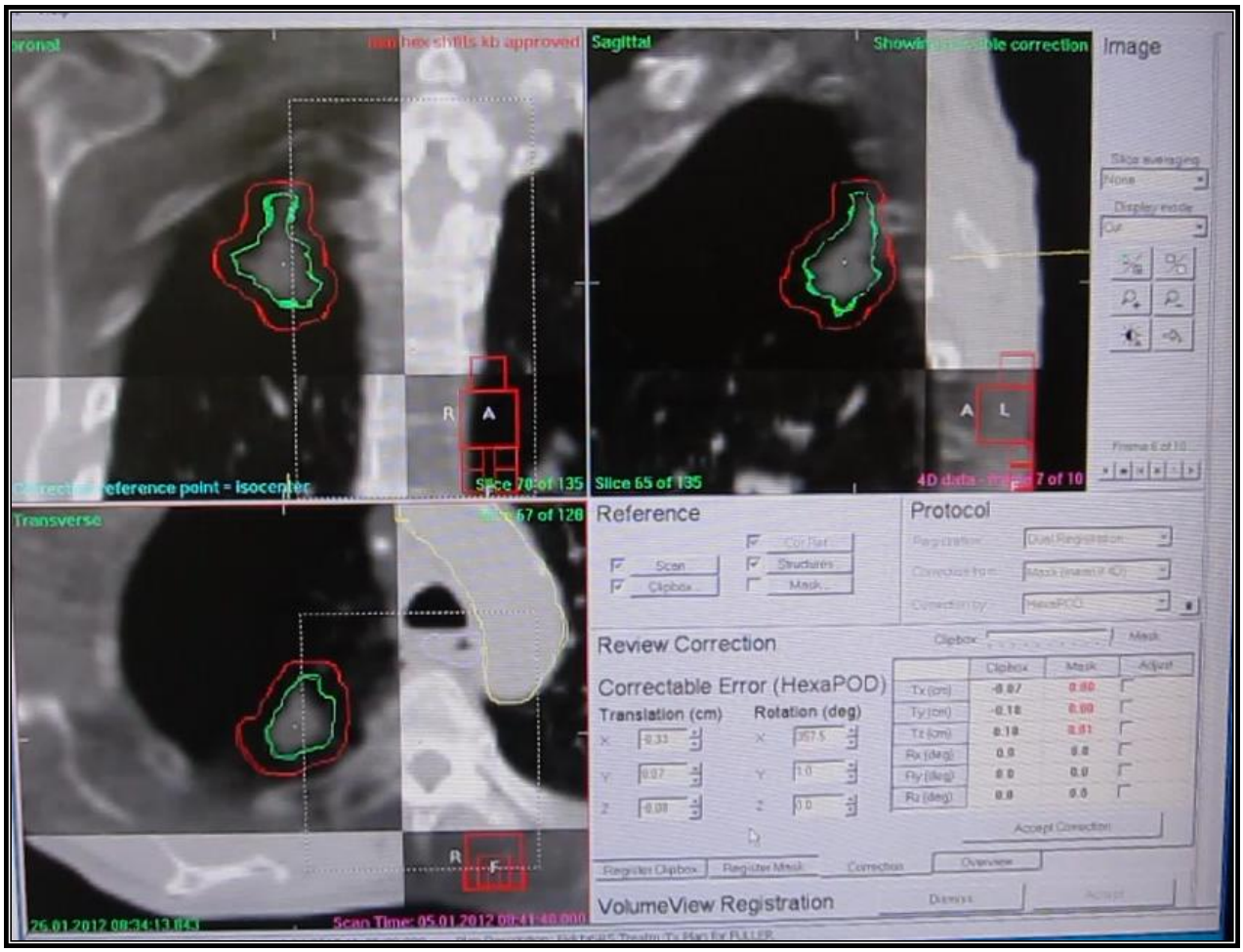


Figure 4 kV CBCT images matched to the reference images to setup the patient to the simulation position

Xing et al. [2006] summarized the recent advancements in IGRT and discussed various practical issues related to the implementation of the new imaging techniques available to radiation oncology community. The key benefit of IGRT technology is that it allows physicians to greatly reduce the added margin since the tumor location can be visualized moments before the treatment is administered. This is an important feature for the following 2 reasons:

1) A reduced margin translates into less normal tissue within the radiation field, suggesting that there is a significant reduction in the risk of damage to the immediately adjacent structures.

2) Reducing the treated area also allows increase in the radiation dose that can be safely administered. Again, the same rules that governed a century ago continue to apply: the higher the dose, the greater will be the chance for cure.

In summary, IGRT allows verification of the target location at the time of treatment. It permits the physician to tailor patient-specific PTV margins. With IGRT, field margin and treatment dose can be routinely customized to each individual patient to achieve a safe dose escalation. Finally, it facilitates adaptation to on-treatment changes by correcting or moderating setup changes, assessing anatomical changes, or in some cases, re-planning. In other words, adaptive therapy becomes possible.

2.1.5 Adaptive Therapy

Yan et al. [1997] describe adaptive radiation therapy as "a closed-loop radiation treatment process where the treatment plan can be modified using a systematic feedback of measurements." The idea of adaptive radiation therapy is to systematically monitor treatment changes and incorporate these variations to re-optimize the treatment plan during the course of treatment. There are many Adaptive Radiation Therapy (ART) techniques that have been developed to exploit the additional information provided by the imaging devices [Martinez et al., 2000; Keller et al., 2003; Ferris and Voelker, 2004; Harris, 2009].

There are two approaches to ART – on-line or offline. In on-line ART, the treatment is adapted to variations during the delivery of a fraction, while the patient is lying on the table. In contrast, off-line ART waits until after the treatment is complete and then retrospectively uses the measurements from imaging devices to make adjustments. Since in offline ART, the imaging data does not become available until after the delivery of a fraction, it allows for a more thorough

and careful analysis. Another advantage of off-line ART is that it does not affect the duration of a fraction, and therefore the patient throughput, since the treatment is not modified during the delivery. In the remainder of this dissertation, whenever the term ART is used, it can be assumed to mean off-line ART. The next section describes tomotherapy, an IGRT treatment device featured prominently in this dissertation, with imaging capabilities well-suited for ART.

2.1.6 HELICAL TOMOTHERAPY

The Tomotherapy Hi-Art System is an IGRT delivery system developed by TomoTherapy Inc. in Madison, WI. The technology was pioneered mainly by professor Thomas Mackie, PhD, students Timothy Holmes and Stewart Swerdloff, and researcher Paul Reckwerdt [Mackie et al. 1993]. The first commercial version, termed the HI-ART 2, was clinically implemented at the Thompson Cancer Survival Center, Knoxville, TN, USA, in July 2003 [Mackie 2006]. The AAPM recently published Task Group 48, "QA for helical tomotherapy" that offers a comprehensive review of the tomotherapy technology and provides recommendations for the quality assurance that should be performed when using this modality [Langen et al. 2010].

The name "Tomotherapy" literally means slice therapy [Mackie et al. 1995]. The name signifies that the radiation is delivered "slice by slice" because of the helical nature of the system. The tomotherapy unit essentially consists of a linear accelerator mounted on a CT slip-ring gantry. The patient lies on the treatment couch, which moves into the gantry as a narrow 6 MV radiation beam continuously rotates around the patient. The couch moving at the same time the gantry is rotating, resulting in a helical pattern radiation beam around the patient, which targets tumors with optimal levels of radiation. This method offers the advantage of having a radiation treatment beam projected into the tumor continuously as it rotates rather than having a limited

number of fixed beams, each providing only a fraction of the dose necessary to irradiate the tumor. Figure 5 presents a diagram illustrating the tomotherapy system.

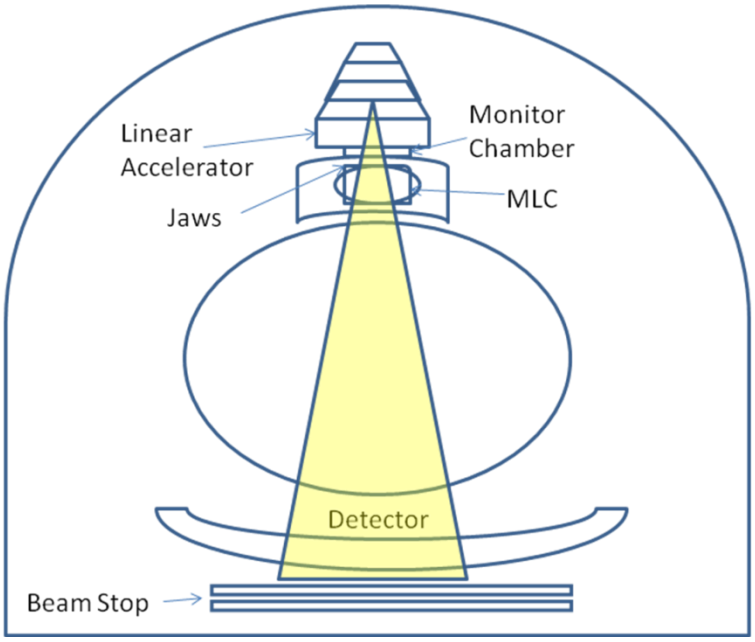


Figure 5 Basic schematic of the tomotherapy Hi-Art System

A unique aspect of helical tomotherapy is that radiation is delivered with a fan beam rather than a traditional cone beam. The thickness of the radiation field from this fan beam is reduced as the jaw setting is selected as 1.0, 2.5, or 5 cm during the planning stage. However, the width of the beam is maintained up to a maximum of 40 cm. The system also has a beam stopper on the rotating gantry opposite the accelerator. The beam stopper consists of 13-cm-thick lead slabs that act as a counterweight and primary beam attenuator [Ramsey et al. 2006].

IMRT delivery is achieved by moving the 64 individual leaves into and out of the fan beam. The MLC leaves are made of tungsten and have a width of 6.25 mm at the isocenter. The MLCs are pneumatically driven and considered binary. The term “binary” represents the 2 discrete positions allowed for the collimator leaves: open or close. The length of time that a leaf is open

is proportional to the intensity of radiation allowed through at a particular portion of the beam. The MLC leaves are designed with an interlocking tongue-and-groove design to minimize leakage when adjacent leaves are closed.

Other than the jaw setting, the pitch and modulation factor are other factors impacting the tomotherapy delivery that must be selected during the planning process. The pitch represents the distance that the couch moves into the gantry bore for each rotation of the gantry. For example, a pitch value of 1/3 will result in the couch moving into the bore at 1 beam width for every 3 rotations. A smaller pitch value would allow for steeper dose gradients but also result in longer treatment times. The modulation factor represents the ratio of maximum leaf opening time to the mean leaf opening time of all MLC leaves open in a given projection [Kissick et al. 2005]. As opposed to the pitch value, increasing the modulation factor will result in longer treatment times but aid in reducing the dose to unspecified tissue and sensitive structures.

Tomotherapy employs an arc-shaped CT xenon detector to acquire the pretreatment MVCT images. The high quantum efficiency of the detector results in improved image quality [Hinderer 2003]. The images obtained before the treatment are used to correct for the positional uncertainties associated with setup error and inter-fraction organ motion. The position of the target volume relative to the treatment isocenter can then be corrected by moving the patient with appropriate offsets.

The tomotherapy detector consists of 738 detector cells. Each cell is comprised of 2 gas cavities divided by a thin tungsten septal plate. The distance between the 2 plates, 0.32 mm, defines the

size of a single gas cavity; the thickness of the plates is also 0.32 mm. The septal plates are 2.54 cm long in the beam direction. A potential of 1300 V is applied across every odd plate, and the even plates function as charge-collecting electrodes for the charge produced in the gas cavities. The gas cavities are filled with xenon gas under high pressure. The front face of the detector is placed 129.2 cm away from the photon source. Since the radius of the detector arc measures only 103.6 cm, the geometrical center of the detector is slightly off-focus with the source. The detector focus point is in front of the detector at a distance of 25.6 cm away from the photon source, directed towards the isocenter. In addition to the xenon detector array, there are 3 ionization chambers located between the rotating X-ray target and the multi-leaf collimator. These sealed ionization chambers measure the machine output prior to attenuation by the patient and store the data in a binary file called a sinogram. This is a binary file that contains data for each projection. There are several types of sinograms in tomotherapy, such as imaging sinograms derived from detector data or control sinograms that contain fluence or MLC data for each projection [Langen et al. 2010]. The sinograms referenced in this work are those collected by the detector array during treatment. These sinograms record the radiation exiting the multi-leaf collimator and passing through the patient during the treatment. Chapter 4 describes these sinograms in greater detail, as they are an integral part of this research.

Overall, this work hinges on the tomotherapy detector array. Chapters 3 and 4 describe how the information obtained by the detector (both in the form of the MVCT images and in the treatment sinograms) contain data that have not been previously used to advance radiation therapy treatment. Both the daily MVCT images and the treatment sinograms can be fed into predictive models to determine information about both patient response and machine performance. Chapter

3 discusses how information thus obtained from the daily MVCT scans taken before treatment initiation can be used to adjust the patient as well as measure the tumor response. Chapter 4 will detail the applications of empirical nonparametric models for detecting errors in the compressed tomotherapy sinogram data. Both these applications illustrate how predictive nonparametric models can be effectively applied to radiotherapy data. These types of models are being widely adopted in other industries, particularly finance- and engineering-driven fields requiring the estimation of performance levels of a system. However, although these models are specifically suited to resolve many of the problems in radiation therapy, they have not yet been widely applied in this field. The next section explains this concept with a brief overview of nonparametric regression modeling.

2.2 NONPARAMETRIC MODELING

The definition of modeling may vary depending on the application, but the basic concept remains the same. As noted by Atkeson et al. [1997], modeling can be defined as "the process of solving physical problems by appropriate simplification of reality." Similarly, Sorakin [2005] defines a model "as a description of a system designed to help an observer to understand how it works and to predict its behavior." From this definition, models are then categorized as either mechanistic or empirical. It is important to note that there are many ways models can be classified, and few statisticians or researchers agree on the classes or subclasses that fall under the general and very broad term of "modeling." However, Figure 6 presents the modeling hierarchy that is adopted in this work.

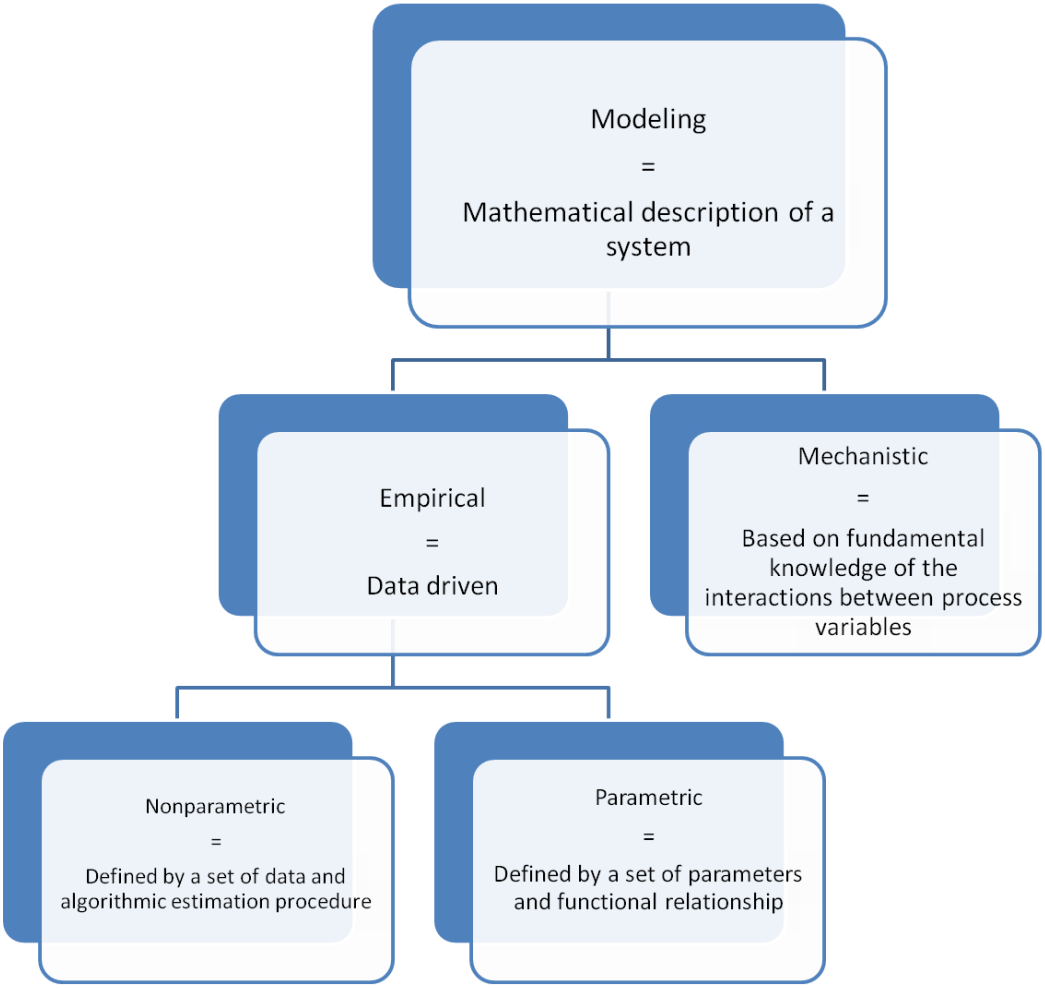


Figure 6 Modeling hierarchy explained

Mechanistic models have a structure that explicitly represents an understanding of biological, chemical, and/or physical processes. A huge benefit of mechanistic models is that they do not require much data for development, and are therefore not subject to peculiarities in data. However, because they require a fundamental understanding of the physics and chemistry governing the process, they can often be very time consuming and occasionally impossible to construct in cases where the interactions between the process variables have not been fully understood. Velten [2009] offers a thorough primer on mechanistic modeling that highlights their many advantages and applications, especially in biology and thermodynamics. In contrast

to mechanistic models, empirical models are solely based on the observations found in the collected data. Empirical models are not derived from assumptions concerning the relationship between variables and are not based on physical principles. Empirical modeling has become a widely used analytical tool in almost all scientific fields. It has gained popularity in recent years, most likely, due to the fact that this type of modeling can still produce accurate and reliable results without requiring intimate knowledge of the underlying (and usually very complicated) first-principals which govern the system being modeled. This work focuses on previously mentioned empirical models, specifically nonparametric empirical models. To understand nonparametric empirical modeling, let us first consider parametric models.

There are really only 3 steps to obtain predictions using a parametric model. First, the parameters of the model must be estimated. The parameters are usually obtained using a set of training data, which have known predictor and response variables and are believed to be un-faulted. The model parameters are optimized to best fit the training data and the data are discarded. The state is then estimated. Finally, the given observations are plugged into the model to predict for the unknown response variable or the future results.

To illustrate this, consider the following polynomial model:

$$y = b_0 + b_1\, x_1 + b_2\, x_2 + b_3\, x_1\, x_2 + b_4\, x_1^2 + b_5\, x_2^2$$

In order to completely define this model for a given set of training observations, the polynomial coefficients b_i are optimized to minimize an objective function, usually the sum of the squared error (SSE). Once the optimal polynomial coefficients have been estimated, the model is completely specified by the equation and the estimated coefficients. The data used to develop

the model are no longer needed and are not retained. Therefore, a parametric model may be roughly defined as a model that may be completely specified by a set of parameters and a functional relationship for applying these parameters to new data in order to estimate the response.

In contrast, a nonparametric model stores historical data exemplars in memory and processes them when a new query is made. For instance, rather than modeling a whole input space with a parametric model such as a neural network or linear regression, local nonparametric techniques may be used to construct a local model in the immediate region of the query. These models are constructed "on the fly" and not beforehand. When the query is made, the algorithm locates historical exemplars in its vicinity and performs a weighted regression with nearby observations. The observations are weighted with respect to their proximity to the query point. In order to construct a robust local model, one must define a distance function to measure what is considered local to the query, implement locally weighted regression, and in some cases consider additional regularization techniques. The clear advantage of a nonparametric technique is that the model structure is not specified apriori but is instead determined from data. Rather than determining parameters (such as the regression coefficients in linear regression or weights in biases in neural networks) and then discarding the historical data, nonparametric models retain all data and use them in the predictions. Nonparametric models decide which observations in the historical data are representative of the current query, and only use observations which are similar to the query to make the prediction. This local characteristic of nonparametric modeling is especially desirable for modeling radiation therapy data. When undergoing radiation treatment, every patient is unique and not all patients will respond the same, meaning global

models are not always appropriate. However, there are definite trends in patient response - meaning local models that can dampen out unnecessary data ideal for these applications. The term nonparametric is not meant to imply that such models completely lack parameters but that the number and nature of the parameters are flexible and not fixed.

There are several factors that need to be considered before deciding whether a parametric or nonparametric model should be used. Both models have clear advantages and disadvantages. In fact, the "no free lunch" theory states that no modeling technique is better than all other techniques on all problem types [Wolpert 1997]. Some of the main benefits and drawbacks of the different modeling types are listed in Table 1. Recent work by Hines and Garvey [2006] discusses the pros and cons of the different model architectures in greater detail and also addresses how one of the main drawbacks of nonparametric modeling can be overcome by using robust vector selection methods to improve the quality of trained locally weighted regression models.

Table 1. Nonparametric vs. parametric models: Key advantages and drawbacks of each model

NONPARAMETRIC	PARAMETRIC
• Can continuously adapt to additional operating regions because new observations may be easily added to the memory vector • Can easily compute uncertainty in the model's predictions with straightforward analytic equations • Has longer recall time, so large models may be difficult to implement in real-time systems	• Requires time-consuming and often unstable updating or retraining to expand to new operating regions • The uncertainty of many neural network predictions cannot be computed without using time-consuming Monte Carlo techniques • Has a more straightforward architecture and shorter run-time

Since the early 1990s, computer and software performance has reached a level of sophistication compatible with nonparametric modeling. Computers can now handle large databases of memory vectors. The data used to develop models can be retained without exceeding the capacity of the computing system. As such, many industries have embraced nonparametric modeling as a practical tool for prediction, diagnostics, and even prognostics.

Garvey's PhD dissertation [2007] shows how these nonparametric techniques can be used to detect, diagnose, and prognose faults and failures in the hydraulic steering system of a deep oil exploration drill. In finance, similar nonparametric models have been employed for options pricing and prediction of high-frequency transactions yield curve data [Lai and Xing 2008]. The NUREG-CR 6895 series explains how the nuclear industry is implementing these techniques to monitor nuclear plants and extend the required interval for sensor recalibration [Hines et al.

2006, 2007]. Similarly, Hubberstone's dissertation [2010] describes an Adaptive Non-Parametric Model (ANPM) that has been developed for integrated monitoring, diagnostic, and prognostic use on small to medium size nuclear reactors. In biology, nonparametric models were used to quantify the effects of wind direction and wind speed on background NO_2 concentrations [Donnelly et al. 2011]. There are so many different and varied applications of nonparametric modeling in almost every scientific field that it is impractical to highlight more than a few. As Fox [2010] summarizes, "Unthinkable only a few years ago, methods of nonparametric analysis and prediction have been rendered practical by advances in statistics and computing, and are now a serious alternative to more traditional parametric methods."

Although predominate in almost every other field, these nonparametric predictive models have not been widely used in radiation therapy. However, their parametric counterparts are used for virtually every aspect of radiation therapy treatment planning and delivery. Most treatment planning systems employ empirical models to calculate dose. Most of these models generalize from measurements in specific conditions to predictions of dose in patient. Korhonen's dissertation [2009] provides a history of the models used by different treatment planning systems. Empirical models are also use to determine EPID response for the purpose of pre-treatment IMRT dose verification [Khan 2008].

Furthermore, empirical models are the cornerstone of radiation biology. These radiobiological models are used in modern radiotherapy to evaluate the biological effects of different treatment plans or modalities. A radiobiological model typically converts a physical quantity (e.g. absorbed dose) to a biological quantity (e.g. cell survival fraction). The two most common empirical

models are the Probit and Logit models. Both models derive mathematical expressions to describe by the probability of an effect following radiation dose. For a more thorough review of radiobiological models, the reader is referred to the article entitled "The progress of radiobiological models in modern radiotherapy with emphasis on the uncertainty issue" by Wang [2010].

Artificial neural networks, which are parametric models inspired by the structure of biological neural networks, have also found wide application in radition therapy. Leszcynski et al. used an artificial neural network (ANN) to evaluate treatment set-up plans. In that study, an ANN was used to accept or reject the radiation dose-volume histogram data for the rectum and urinary bladder. In another study, Kaspari et al. used ANNs to predict desired coefficients for polynomials to describe PTVs for glioblastoma multiforme tumors. These are only two of the many examples of how neural networks have been used to improve radiotherapy treatments. The book *Intelligent and Adaptive Systems in Medicine* presents a more extensive review of neural networks in radition therapy [Haas 2008].

In short, parametric models are entirely engrained in radiation therapy. They are a driving factor of the field. Without the use of parametric models, it would be nearly impossible to treat cancer patients with today's technology. However, nonparametric models are not as prevalent. There are very few instances in the literature of nonparametric models being used in the field. Seibert [2007] and researchers at Stanford [Li and Xing, 2011] have independently investigated using nonparametric techniques to predict the respiratory motion. Additionally, researchers at Washington University have considered using nonparametric methods to learn the complex

interactions between observed toxicities and treatment [Bradley and Deasy, 2008]. Bowling [2012] also has employed nonparametric models in his work developing Gamma Knife plan optimization techniques. In Bowling's research, the model is not used for regression, but is actually a selection technique. He uses a nonparametric model to select a set of plans from past treatment volume that are similar to the current treatment volume. The parameters from these past plans are then used to initialize the Gamma Knife optimization routines that will be applied to the current treatment volume.

Other medical fields that are intertwined with radiation therapy rely on nonparametric techniques. For instance, in medical imaging, nonparametric techniques are used to reconstruct, restore, interpolate, optimize, and analyze MRI, PET, and CT scans [Dougherty 2009]. But, overall, the examples of nonparametric prediction techniques being used in radiation therapy are few and far between. In general, it seems as if the field of radiation therapy has been reticent in implementing these novel methods. This is unfortunate, as nonparametric techniques can be powerful tools that have the potential to be very beneficial. The next sections will describe how these modeling techniques are fundamental in the development and implementation of PART.

3. PREDICTIVE ADAPTIVE RADIATION THERAPY FOR LUNG TUMORS

This chapter describes the basic methodology and results from the locally weighted regression lung tumor response model. The main purpose of this model is to determine if the final change in lung tumor mass and volume could be predicted using only information collected by the tomotherapy MVCT scans obtained early in the treatment. The prediction of the final tumor mass and volume gives clinicians valuable information about how the tumor is responding to radiotherapy treatment and could even lead to the treatment being adapted to account for these changes. Adaptation of the treatment based on the results of the predictive model is a novel concept termed Predictive Adapative Radiation Therapy, abbreviated PART.

3. 1 INTRODUCTION TO TUMOR RESPONSE MODELING

Recent studies have shown that lung tumors can substantially change in size, shape, density, and center of movement during the course of radiation therapy [Kupelin et al. 2005; Underberg et al. 2006]. These changes can be measured with frequent pre-treatment CT imaging, and potentially provide information on patient outcomes [Kim et al. 2005]. Although the underlying biological mechanisms for tumor responses in individual patients are currently unknown, the response can be measured directly during IGRT using soft-tissue imaging. An assessment of the response to therapy during treatment could be used to help direct the patient's further therapy. Unfortunately, patients typically complete with their course of radiation therapy before the tumor response is completely analyzed. Thus, a nonparametric model that employs a database of past tumor responses to predict the response for new patients during treatment using limited

observations collected early in the treatment. This forecasting of tumor behavior can potentially be utilized to adjust the course of treatment.

Research on image-guided radiation therapy (IGRT) for lung cancers has primarily focused on the application of stereotactic body radiation therapy (SBRT) and respiratory management techniques [Barnes et al. 2001; Underberg et al. 2005]. The development of CT-based IGRT systems has enabled clinicians and lung cancer researchers to volumetrically image internal anatomy prior to the delivery of radiation treatment fractions. These images are primarily used to position the patient prior to treatment based on the known position of the tumor and the location of the surrounding normal tissues, resulting in potential reduction of the patient setup uncertainty and the internal uncertainty of the target position [Scrimger et al. 2003; Kron et al. 2004].

In addition to tumor localization, the CT images acquired during IGRT can be used to measure and evaluate the response to treatment (Figure 7). Recent studies have shown that lung tumors can substantially change in size, shape, density, and center of movement during the course of radiation therapy [Kupelin et al. 2005; Underberg et al. 2006]. These changes can be measured with frequent pre-treatment CT imaging, and can potentially provide information on patient outcomes [Kim et al. 2005].

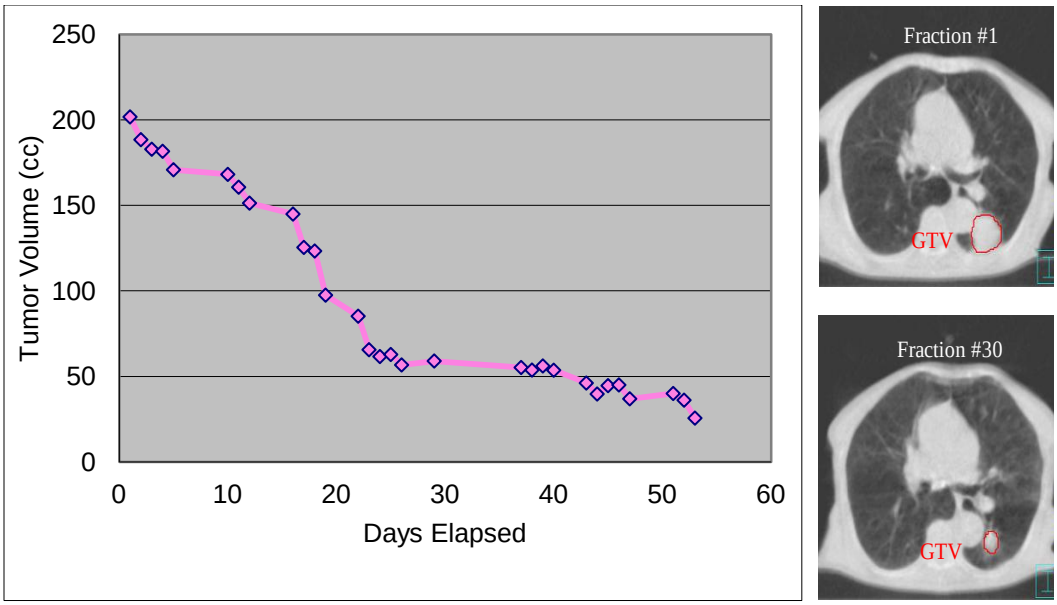


Figure 7 Tumor response measured from daily MVCT scans

Although the underlying biological mechanisms for tumor responses in individual patients are currently unknown, the response can be measured directly during IGRT using soft-tissue imaging. An assessment of the response to therapy during treatment could be used to help direct the patient's further therapy. Unfortunately, patients typically complete their course of radiation before the tumor response is completely analyzed.

In this research, a database of past tumor responses is used to predict the response for new patients during treatment using limited observations collected early in the treatment. This forecasting of tumor behavior can potentially be utilized to adjust the course of treatment. It will allow for implementation of ART. The predictions could also be used in the design of an intensity modulated integrated boost to the location of the residual tumor mass [Dirkx et al. 2004].

3.2 METHODS AND MATERIALS

3.2.1 TUMOR RESPONSE

Tumor responses were measured in 37 patients with lung cancer who visited the Thompson Cancer Survival Center, London Regional Cancer Centre, and the M. D. Anderson Cancer Center Orlando between July 2003 and April 2008 for imaging and treatment with helical tomotherapy (HI-ART, TomoTherapy, Inc., Madison, WI). Several other patients were evaluated, but they primarily had medial disease and were not included because their disease could not be adequately assessed with MVCT imaging. A total of 40 lesions were included in the study. Patients were treated with different doses ranging from 2.0–2.5 Gy per fraction according to different institutional preferences and protocols. The total doses and treatment fields were implemented at the discretion of the radiation oncologist. Table A-I (found in Appendix A) presents the patient demographics of the study participants.

Megavoltage CT (*MVCT*) images were acquired prior to the delivery of the lung treatment fractions for the 37 patients. The primary purpose of the MVCT images was to position the patients for treatment using an automatic CT-to-CT fusion with the treatment planning CT images. All MVCT images were fused relative to the original treatment planning CT using soft-tissue anatomic landmark-based fusion. The tumor volumes (excluding nodal disease) were outlined on each MVCT slice using the same window and level settings, which were set to enhance the visibility of the lung tumor relative to the surrounding tissue. Fifteen tumors were automatically contoured using Model Based Segmentation of the original tumor volume contours from the original treatment plan. The tumor volumes (and some masses) were calculated by the

treatment planning systems based on the defined contours. The volume and masses were recorded against the elapsed day of treatment for all 40 of the tumors in the study.

Additionally, some tumors were re-contoured by different investigators to eliminate inter-individual variation. The variation in the contouring between 2 investigators was found to be less than 2%. This low inter-observer variability may be explained by the use of the original treatment planning CT contours as a baseline and that 19 out 20 lesions were surrounded almost entirely by lung without atelectasis. Systematic inter-observer errors were present, with one investigator consistently contouring smaller tumor volumes. However, the absolute change in tumor volume remained relatively constant regardless of the investigator performing the contouring.

The data were relatively complete with a mean of 29 CT image sets per patient. Missing data were replaced using a piecewise cubic spline interpolation function [Press et al. 1992]. All replaced data were flagged. The reconstructed data were then used to develop models to predict tumor response, and the models were created using as little of the replaced data as possible.

3.2.2 TUMOR RESPONSE MODELING

The tumor responses measured in this study varied greatly from patient to patient. Figure 8 shows the change in tumor volume (normalized to the first day of treatment) versus the elapsed days of treatment for several tumors. These tumor responses vary greatly and cannot be modeled with a simple linear or low-order polynomial fit. Rather, a more sophisticated modeling technique that can learn complex relationships was required. The linear weighted regression (LWR) model was initially chosen to apply to this data because of its ease of retraining or

learning new cases, its ability to determine if the current patient regression is similar to that of historical patients, and because analytical techniques are available for uncertainty quantification.

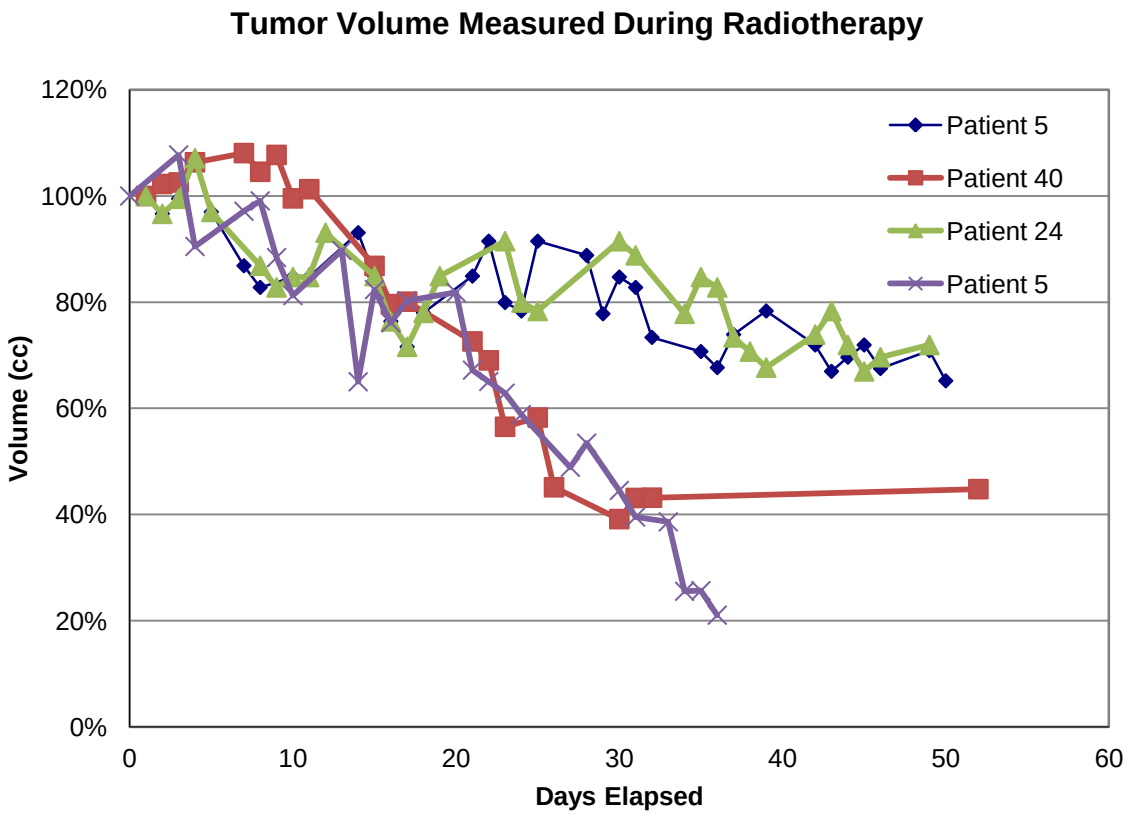


Figure 8 Tumor Response is considerably more varied for a large cohort of patients

Locally weighted regression (LWR) is a nonparametric memory-based modeling technique. LWR stores the training data in memory and recalls the relevant data exemplars when a query is made. When a query is processed, the algorithm locates similar training exemplars and performs a weighted regression with the nearby observations [Atkeson et al. 1996]. This is different from parametric techniques such as neural networks that use an optimization algorithm to iteratively calculate the model used to represent a desired input-output relationship. Parametric training procedures are subject to local minima when non-linear techniques are employed, which may result in poor models. Training very complex relationships needed for regression prediction can

take days and optimizing the architecture to provide robust results requires extremely experienced developers. Additionally, if new data are acquired, the model must be retrained requiring a significant effort. Nonparametric techniques such as the one proposed in this research require no training and new data can easily be added to the historical memory matrix.

LWR is a non-linear, nonparametric extension of linear regression. It uses a weighted sum of squared errors, where the weighting factor considering the distance of the training point from the query point is minimized, to give more training points that are closer to the query point. The notation $Q(q)$ is used because the function is optimized around query point q [Fan and Gijbels 1996].

$$Q(q) = \sum_{i}^{n} (y_i - x_i w)^2 K(d(x_i, q)) \qquad (3\text{-}1)$$

where: x_i are the training points

q is the query point

d is a distance function such as the Euclidean distance

K is a kernel weighting function.

A typical kernel weighting function is the Gaussian kernel with a vector of local bandwidths (h) [Wand and Jones 1995]:

$$K(d) = \frac{1}{\sqrt{2\pi h^2}} e^{-d^2/2h^2} \quad , \qquad (3\text{-}2)$$

with the Euclidean distance given as

$$d(x,q) = \sqrt{\sum_j (x_j - q_j)^2} \quad , \tag{3-3}$$

where: j is the dimension of the vector.

The bandwidth determines the definition of local, i.e., how close a training point needs to be to a query point to allow it to influence the prediction. All the training points within the specified bandwidth are used to calculate the prediction for a given query case. However, the training points are entered into the prediction with different weights. Training points that are nearer to the test point are given more importance. Nevertheless, the bandwidth controls the smoothness of the predictions by deciding which neighboring training points will be considered in the prediction. When the bandwidth is small, the local linear estimator interpolates the data points and over-parameterizes the unknown function resulting in noisy estimates. When the bandwidth is large, the local linear estimator reduces to the standard parametric linear regression estimate, which, with the exception of the cases where the model is correctly specified, under-parameterizes the regression function, resulting in a large modeling bias.

This study employs local bandwidths, rather than a single global bandwidth. Local bandwidths allows each input variable to have its own unique corresponding bandwidth value; this has been shown to improve model performance [Gerard and Schucany 1999].

The following equation solves the above weighted least squares regression equation for the optimal estimates of the regression coefficients:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{W}^T\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}^T\mathbf{W}\mathbf{Y} \tag{3-4}$$

where W is a diagonal matrix with the diagonals being equal to the square roots of the kernel function:

$$w_{ii} = \sqrt{K(d(x_i, q))} \tag{3-5}$$

Solving the regression equation results in a new estimator:

$$\hat{y}(q) = q\left(\mathbf{X^T W^T W X}\right)^{-1}\mathbf{X^T W^T W Y} \ . \tag{3-6}$$

Implementing LWR required first selecting the training data (past tumor volume data over the number of treatment days elapsed). Volume data from the 40 historical lesions were used to build the initial memory matrix. This large amount of initial data was sufficient to construct a memory matrix capable of producing reliable predictions.

A genetic algorithm (GA) was used to determine the minimum number of test observations that yield an accurate prediction. Since contouring the tumor on MVCT scans is a time-consuming process, it is desirable to contour as few scans as possible. Additionally, an accurate prediction is needed as early in the treatment as possible so that there is ample time to modify the treatment if needed. Using a GA proved to be the most efficient and effective way to accomplish this result.

A GA starts with a set of solutions represented by chromosomes, called a population, with the performance of each solution evaluated by a fitness function [Haupt and Haupt 1998]. In this case, each solution contained the different days from which tumor volume data should be collected and the vector of local kernel bandwidths. For example, one solution might evaluate

the model using volume data from treatment days 1, 3, 5, 7, 10, and 14 and a vector of bandwidths. Constraints were placed upon the allowed number of observations and how late in the treatment these observations could occur. The fitness function evaluated the leave-one-out cross-validation (*LOOCV*) error of the solution. To evaluate the LOOCV error, the model was run for each lesion using the remaining lesion data as training data, and the average absolute error of all predictions was computed. Obviously, there was a trade-off between the model's accuracy and limiting the number of observations or observations that occurred later in the treatment. However it was found that treatment days 2, 4, 5, 9, 11, 12, 15, and 17 elicited the best performance for the smallest number of observations earliest in the treatment. Since the volume data covers several orders of magnitude, the input and output data were mean-centered and scaled to unit variance prior to presentation, and the results were scaled back to their original units.

3.2.3 TUMOR RESPONSE MODELING UNCERTAINTY

Because this technique will eventually be used to make decisions about a patient's treatment, each prediction must be accompanied by a measure of its uncertainty. A large uncertainty value indicates that the prediction is unreliable. In this study, uncertainty is expressed in terms of a confidence interval. The confidence interval (CI) is a quantification of the agreement of model predictions with the system's true value. It provides a measure of the uncertainty of a model's expected prediction. For a confidence level of 1-α, the confidence interval was defined by Tamhane and Dunlop [2001] as the interval [*L*, *U*] including $f(\underline{\hat{x}})$ (the true system value) with a pre-assigned probability of 1-α can be expressed as follows:

$$P[L \le f(\underline{\hat{x}}) \le U] \ge 1-\alpha \ . \qquad (3\text{-}7)$$

For a confidence level of $1-\alpha$, the CI may be written as follows in terms of the variance and bias:

$$E(\hat{y}) \pm t_{n-p,\alpha/2} \sqrt{Var(\hat{y}) + Bias^2}, \quad (3\text{-}8)$$

where: n is the number of training observations

p is the number of variables used to infer y

$t_{n-p,\alpha/2}$ is the t-statistic which approximates the normal distribution for $n-p$, degrees of freedom, and confidence level $1-\alpha$

$E(\hat{y})$ is the expected model prediction of y.

In this equation, the variance measures the random error of the model's predictions, while the bias measures any systematic error. For an LWR model, the variance of the estimator is calculated using Monte Carlo techniques or with the following analytic equation [Hines 2005]:

$$Var\{\hat{y}(q) \mid X_1,..., X_n\} = q(\mathbf{X}_x\mathbf{W}_x\mathbf{X}_x)^{-1}\mathbf{X}_x^T\mathbf{W}_x\mathbf{V}\mathbf{W}_x\mathbf{X}_x(\mathbf{X}_x\mathbf{W}_x\mathbf{X}_x)^{-1}q^T, \quad (3\text{-}9)$$

where: $\mathbf{V} = diag\{v(X_1),...,v(X_n)\}$, and $v(X_i) = \text{var}(Y \mid X = X_1)$

q is the $(d+1)\times 1$ vector having 1 in the first element and 0 elsewhere

d is the dimension of the predictor variable set

This definition applies to local linear modeling only for local constant modeling $p=0$ and $q=1$. The estimator bias is assumed to be negligible. This assumption was validated by comparing the mean squared error (the average squared error of a model's predictions and target values) to the total uncertainty; the results revealed that the variance was the chief contributing component.

3.3 RESULTS

For the 40 easily visualized lesions, the tumor volume response over the treatment course was measured using pre-treatment MVCT imaging (Table A-II). The initial volumes ranged from 1.9–737.2 cm^3, with a mean volume of 102.5 cm^3 and a median volume of 32.6 cm^3. The predicted final tumor volumes ranged from 0.9–375.8 cm^3 with mean and median volumes of 55.2 and 15.1 cm^3, respectively. Tumor regression was correctly predicted in all 40 lesions (Figure 9). Only 15 of the lesions also had mass data, so these predictions are not shown. The mean absolute difference between the true and predicted final volumes was 3.4 cm^3 (8.4%). In the 2007 publication, the mean absolute error with only 20 lesions was 12%, so the additional data did improve model performance. It is also important to recognize that since not all of the patients were treated with the same fractionation scheme, the day that the measurement for the predicted final volume will vary between patients (for instance, some patients ended treatment on day 35, well others were still receiving treatment at elapsed day 50). So for the predictions shown in the Figure 9, the model was given each lesion's final volume as its y-training set.

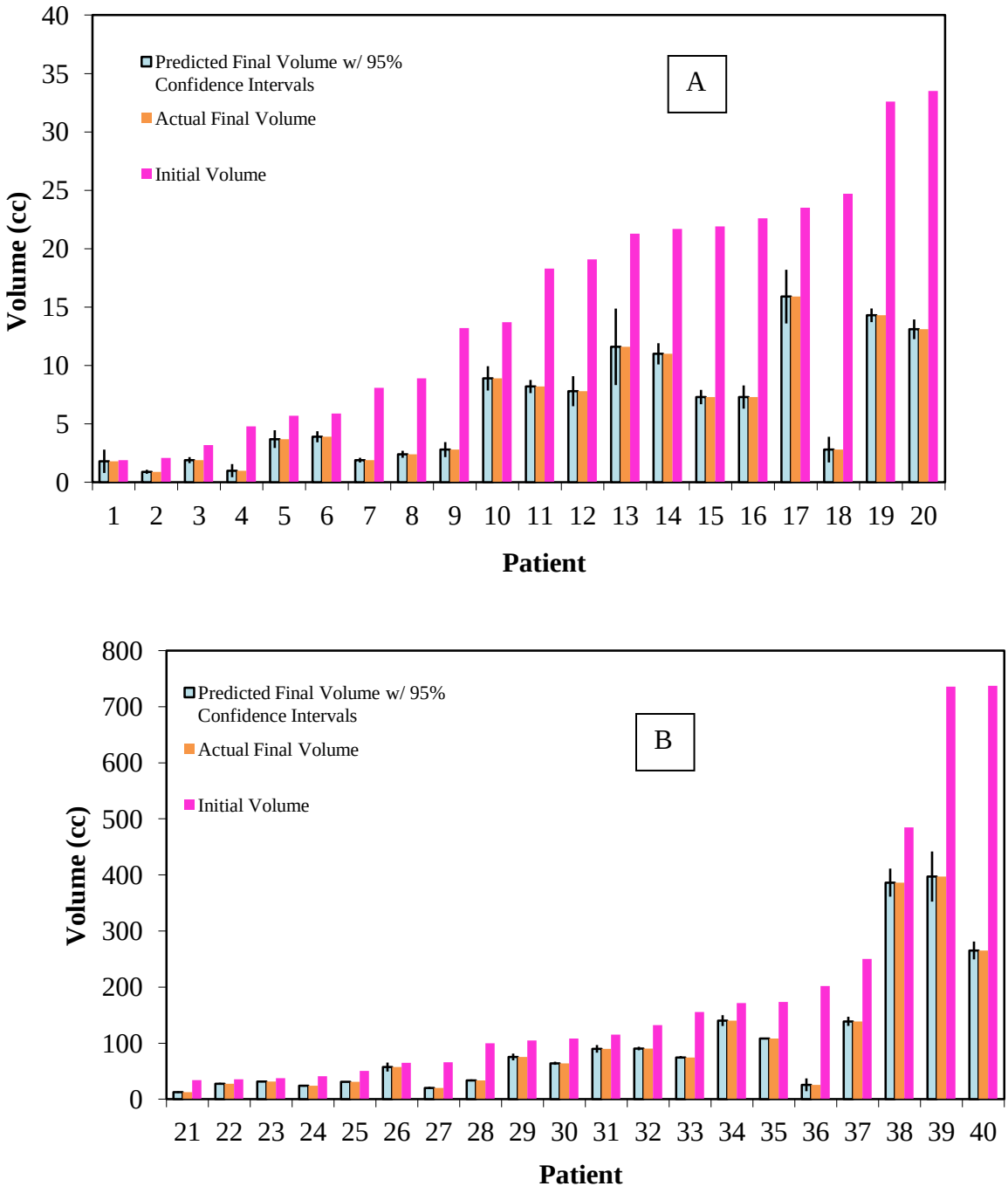


Figure 9 LWR Final Volume Predictions for (A) Patients 1–20 and (B) Patients 20–40

Figure 10 plots the results of the volume and mass models for 3 typical lesions, with the training data (observations used to make the predictions) highlighted by pink circles and the query data (data being predicted) shown in red, surrounded by the 95% error bars.

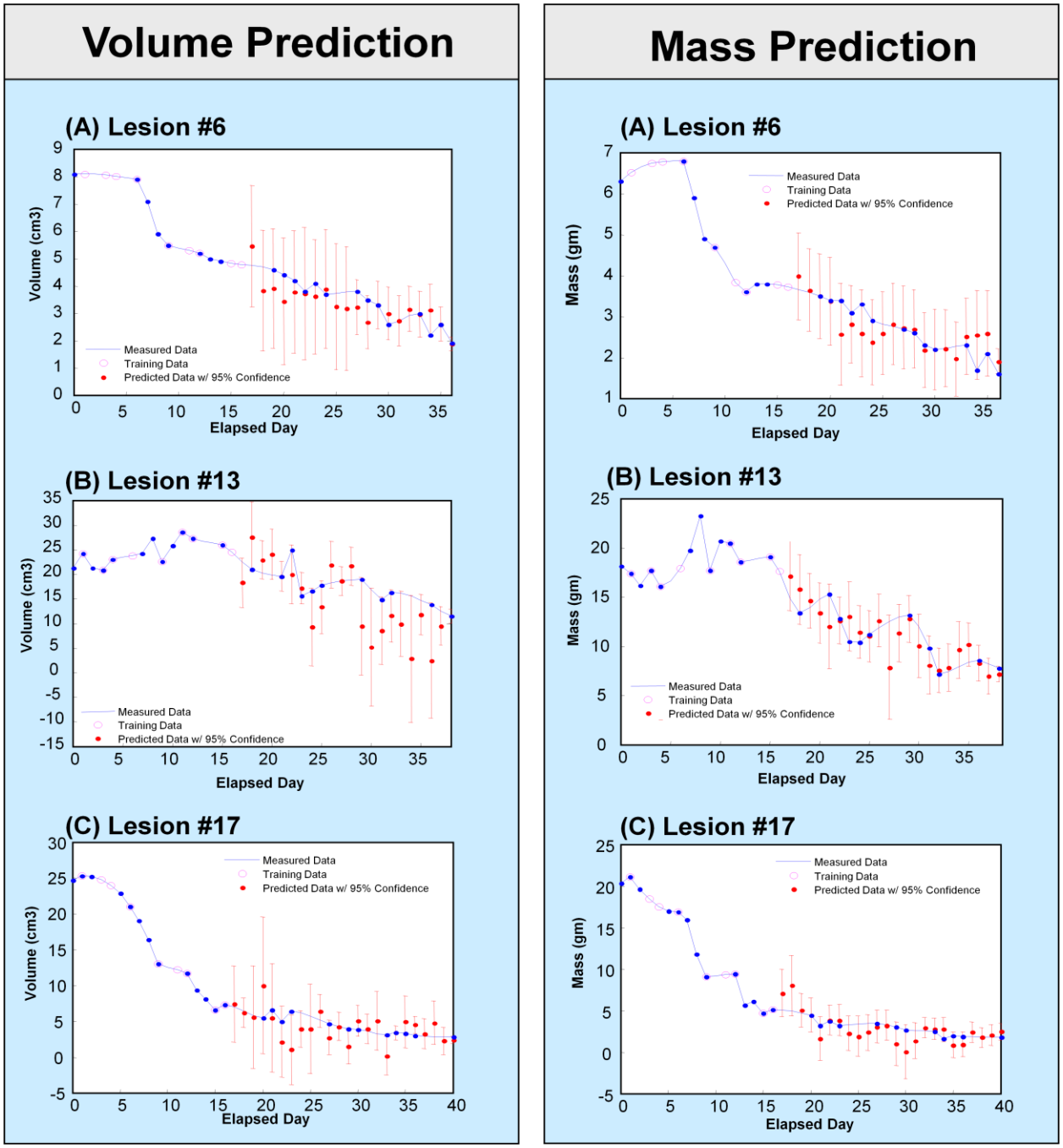


Figure 10 Locally weighted regression volume and mass predictions over the course of treatment for lesions 6, 13, and 17.

All predictions were based on measured tumor volumes and masses from treatment days 2, 4, 5, 9, 11, 12, 15, and 17. This combination of observation days was determined using genetic algorithm-based optimization to minimize LOOCV error and the number of observations. Although the model is extremely accurate in its predictions made at the end of a patient's treatment, the uncertainty in the shape of the tumor response curve increased near the middle of treatment.

Figure 11 plots the average uncertainty of the predictions for all the lesions versus the day the prediction was made for. The graph starts at day 18, since day 17 was the last day the model was allowed to have data from. Since most tumors only grow a small amount from one day to the next, the model can correctly predict the change in volume for day 18, and the uncertainty is low. However, after only several days the uncertainty begins to increase. The graph illustrates that the tumor response relationships themselves increase in complexity toward the middle of the treatment course, and the predictors are more varied, which increases the uncertainty. Respiration motion and contouring variations can result in fraction-to-fraction variations in the measured tumor volumes; these may not represent actual changes in volume. It would be interesting to see how much the uncertainty during the middle of treatment could decrease with a larger database of lesions.

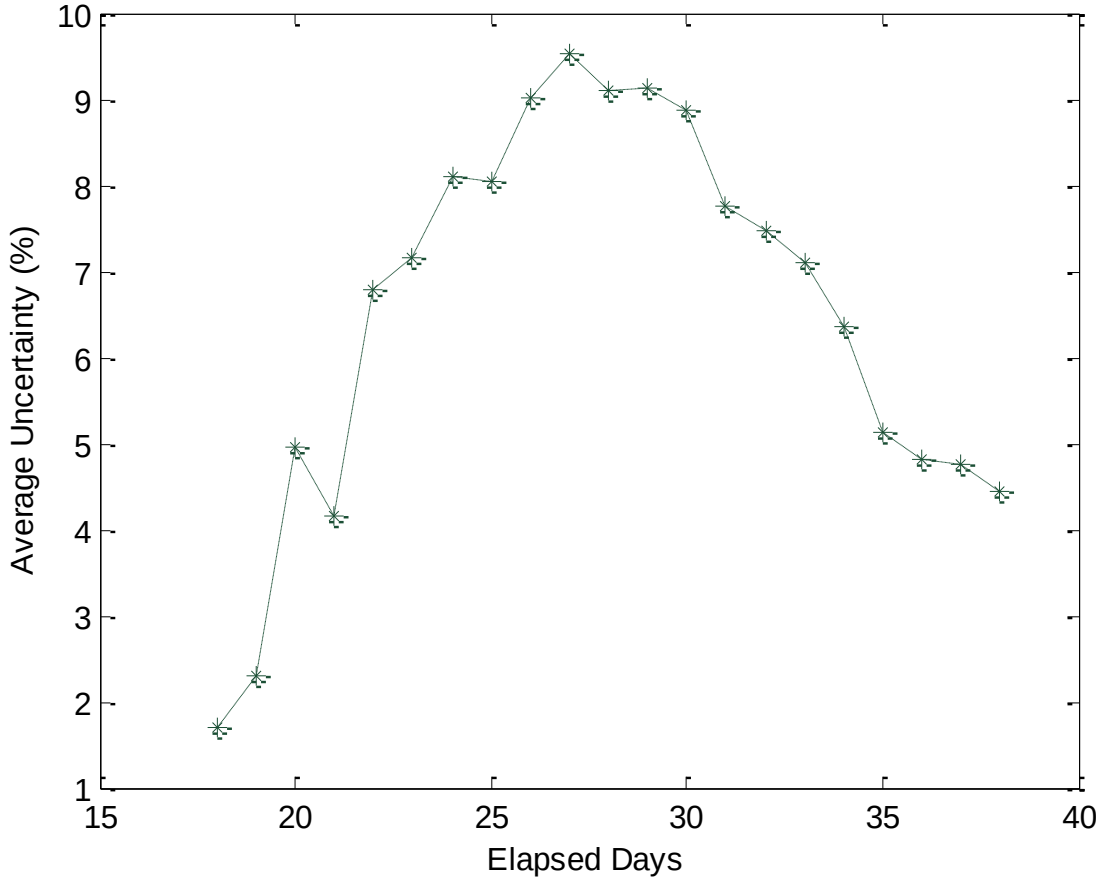


Figure 11 Average Uncertainty for all lesions versus Day of Prediction.

Multiple models were also created that included patient information such as type of tumor and dose. When these variables were included, the root-mean-squared error increased from 15.2% (over prediction for the entire course of treatment) to 20.1%. Since each patient is considered a variable, and the additional patient data were included as observations, it was impossible to use local bandwidth optimization to dampen out unimportant information. This result was surprising, as it was anticipated that this information would be a valuable input. Different trends were observable in the data based on the type of cancer, but the performance of the model did not improved with the addition of different inputs. All combinations with these additional data

points were attempted, along with various scaling methods. Perhaps the LWR technique had trouble handling these categorical variables. This is an area warranting further investigation.

The 95% confidence intervals were analytically calculated for each of the 40 lesions. The confidence intervals are shown in Figures 10 as error bars on the predicted volumes. The confidence intervals are also listed in Table A-II for the LWR predictions of the volume on the final day of treatment. The final volumes with the 95% confidence interval were also predicted correctly in cases where the tumor volumes did not change much over the course of treatment (lesions 1 and 12). However, these predictions have larger confidence intervals. These inflated confidence intervals are expected because the training set had few other examples of this behavior, and calculations showed that the model was being extrapolated outside its training region, thereby increasing the uncertainty.

The predictive performance of the model will increases with the amount of training data. In this study, volume data from the remaining 39 lesions were used for predicting change in volume for each tumor. Because LWR is nonparametric, the memory matrix can be appended with new data. As more patient data is added and the memory matrix grows, the predictive accuracy of the model should improve. The local nature of LWR is also well suited for patient modeling. Since each patient is unique and not all patients will respond the same, global techniques are less appropriate. With local models, only observations which are similar to the current query are used to make the prediction. Since there is varied response amongst the patients, the ability to dampen out observations from patients which are not responding in a similar manner as to the current patient is desirable.

3.4 TUMOR RESPONSE MODELING CONCLUSIONS

A novel technique has been developed for predicting lung tumor response during radiation therapy. This study expanded upon the original by including data from an additional 20 lesions (thus, a total of 40 lesions were included in the memory matrix). Because LWR is nonparametric, the memory matrix was easily appended with this new data without having to retrain the model. The predicted volumes and masses at the end of treatment highly agreed with the true final volumes and masses, and the additional data did improve the model's accuracy. However, inclusion of patient information (such as type of tumor and dose) did not improve the model performance as expected. Overall, this study confirms that predictive modeling, specifically in the context of PART, has the potential to improve radiation therapy. By knowing the expected volume change in a tumor as early as the third fraction of treatment, care can be taken to ensure that the tumor volume remains in the treatment field throughout the entire treatment course. LWR predictions could also be used in the design of an intensity modulated integrated boost to the location of the residual tumor.

Additional studies are needed to determine if there is a correlation between the predicted tumor response during treatment and clinical outcomes. Future research hopes to add PET images acquired before, during, and after treatment into the model and to also correlate patient survival with model predictions. It would also be interesting to extend PART beyond just lung tumors and investigate modeling different anatomical sites.

4. VERIFICATION OF HELICAL TOMOTHERAPY DELIVERY USING AUTO-ASSOCIATIVE KERNEL REGRESSION

Beyond the clinical necessity of ensuring the patient's IMRT plan is correct, billing codes which must be submitted for reimbursement from Medicare and insurance companies also state that patient specific IMRT QA must be completed prior to the patient starting treatment. This means that performing patient specific QA for all new IMRT patients is not just good clinical practice; it also is a regulatory requirement. Unfortunately, performing patient specific QA can be fairly time consuming. Even with the advances in portal dosimetry and other QA measurement tools, it still takes time and resources to deliver the plan and analyze and record the results. Having to re-QA the treatment plan each time it is modified is a major hurdle in implanting PART. Most departments simply do not have the personnel to support constantly reshooting the QA. With that in mind, methods for automating the QA process have been investigated.

This chapter describes the development of a nonparametric modeling technique developed to monitor nuclear plant sensor data can also be applied to model tomotherapy data. Researchers at the University of Tennessee developed auto-associative kernel regression and found that this modeling scheme had excellent performance on the nuclear plant data and could predict correct sensor values when supplied a group of corrupted or faulty sensor values. Although the modeling scheme is young, the promising results from the nuclear plant data, made me wonder if it could also be applied to the field of radiation therapy. After considering the possible applications, I then realized that the exit dosimetry data from the helical tomotherapy system does not differ much from data collected at nuclear plants. Although exit dosimetry data have

considerably higher dimensionality, making the problem more complicated, I still believed the modeling technique could work. The following sections will outline the process of developing the model with the tomotherapy sinogram data and communicate the AAKR results.

4. 1 INTRODUCTION TO AAKR FOR ERROR DETECTION

As previously mentioned, the clinical process for intensity modulated radiation therapy (*IMRT*) often requires quality assurance (QA) testing for individual patients to verify that the correct dose (and dose distribution) is delivered to the correct target site in the body [Molineu *et al* 2005]. This IMRT QA process requires lengthy measurements before each patient's first treatment, and again after any changes (such as re-plans, boosts, and adaptations). When PART becomes available, the IMRT treatment delivery sequence will be adjusted prior to each treatment session to compensate for changes in the position, size, and deformation of internal anatomy. As such, it will be impossible to perform an IMRT QA analysis prior to treatment delivery without moving the patient.

Schemes for verifying the delivery of the correct fluence or dose to the patient using electronic portal imaging devices (EPIDs) have become a subject of considerable interest. Most studies have focused on the application of EPIDs designed for patient setup verification during radiation therapy. The most commonly used systems are video-based EPIDs, scanning liquid-filled EPIDs, and amorphous silicon arrays, and each system has its own advantages and disadvantages. Recently, Dempsey [2006] et al. developed a custom EPID that employs a fast scintillator and a high-speed camera to experimentally characterize static and dynamic IMRT delivery. The vast majority of EPID studies have focused on IMRT verification with either no

material placed in the path of the beam, or with homogonous build-up material placed on the EPID [Louwe et al. 2006]. Few EPID studies have performed IMRT verification with an anthropomorphic phantom in the beam, and even fewer have performed verification for actual treatment deliveries in patients [Warkentin et al. 2003].

In helical tomotherapy treatment delivery, exit dosimetry data are acquired for each patient by a CT detector array located opposite the linear accelerator on a rotating slip-ring gantry. This data is stored in the tomotherapy record and verified automatically for each treatment fraction. Unfortunately, techniques developed for performing IMRT patient QA with EPID detectors cannot be applied to tomotherapy detector data. This work aimed to develop a novel technique for automatically evaluating exit dosimetry on a helical tomotherapy system using auto-associative nonparametric modeling that is robust and has the capability to learn complex detector data relationships, even with low temporal resolution of detector data and patient beam attenuation.

4.2 METHODS AND MATERIALS

As mentioned in Chapter 2, the tomotherapy detector array collects and stores exit dosimetry data during treatment delivery in the form of sinograms. These sinograms contain a record of the radiation exiting the multi-leaf collimator and passing through the patient during the treatment (Figure 12). In the figure, the left panel shows a macroscopic view of the detector data, and the right panel shows an enlarged view of 6 projections. All MLC leafs are closed at the beginning and the end of each projection. The intensity is modulated by adjusting the time that each MLC leaf is open for each projection. The tomotherapy sinograms consist of the expected beam

intensity at the detector array for each gantry angle and couch position. This intensity pattern is referred to as a "sinogram" because each point irradiated in the patient maps a sine wave pattern at the CT detector as the gantry revolves. The horizontal axis in the sinogram represents the X-ray fluence incident on the detector array, while the vertical direction is a function of projection angle.

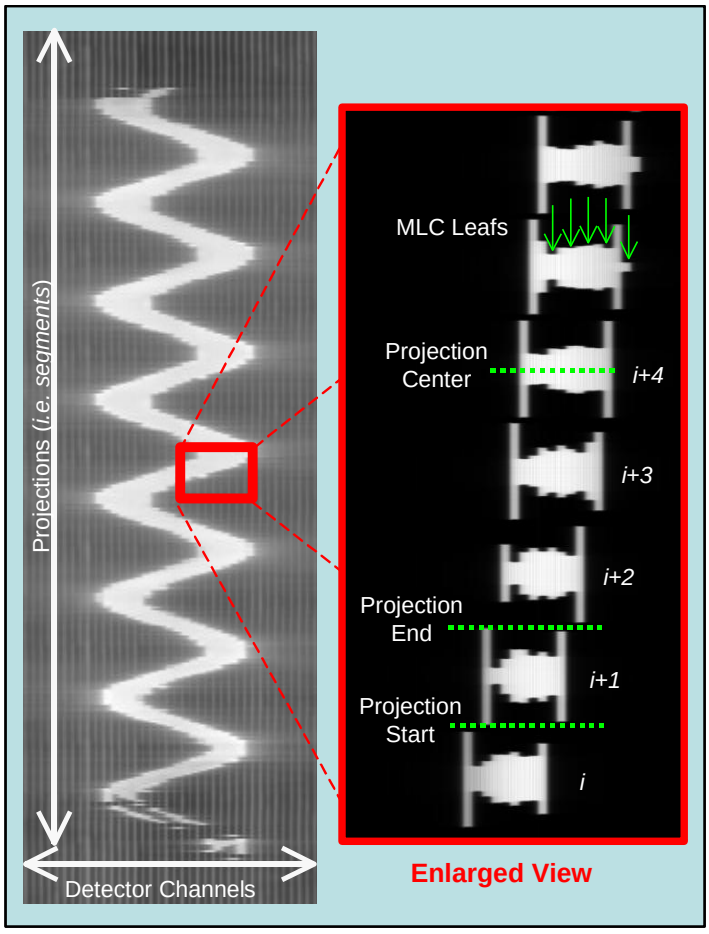


Figure 12 Raw detector sinogram data for helical tomotherapy treatment delivery.

A projection is the planned fluence at a particular gantry angle. The intensity of the sinogram is low for regions of low radiation fluence (i.e., under a closed MLC leaf), while the high intensity values represent regions where radiation is exiting the MLC. This sinogram data can be accessed in uncompressed format (300 frames per second) immediately after treatment delivery from the

tomotherapy data acquisition system (DAS), which is an on-board computer process system located in the gantry. The treatment delivery sinograms are stored in the tomotherapy relational database. The detector data for clinical patients are compressed per projection. There are 51 projections per rotation, and the gantry rotational period ranges from 10–60 seconds. The stored detector data is the cumulative signal acquired during each projection, which ranges from 0.3–1.2 seconds. This detector data is referred to as "compressed" data.

The compressed sinogram exit detector data contains valuable information regarding the integrity of the treatment. Thus, the analysis of compressed sinogram detector data could become a critical component of IGRT quality assurance (QA). Unfortunately, this analysis is not as straightforward as it may seem. Even with accurate setup, a patient's breathing and anatomical shifts will cause the beam's attenuation to slightly vary at each treatment. This change in attenuation is reflected in the compressed sinogram detector data. Because of the difference in attenuation, direct comparison between treatment sinograms using simple sinogram-sinogram subtraction is not always able to discern small errors (such as the incorrect opening of a single MLC leaf). Although sinogram-sinogram subtraction is able to detect gross errors, since the nature of MLC leaf errors in actual treatment delivery is unknown, another method is needed to distinguish positional errors of individual MLC leaves. When analyzing the uncompressed data, geometric shape detection can be used. Chen et al. [2012] describe how a Richardson-Lucy deconvolution can be used to help remove some of the scattering and penumbra that "pollute" the uncompressed sinogram data to make the geometric detection more meaningful.

Unfortunately, even the Chen et al. [2012] study admits that the majority of uncompressed raw data poses a workflow issue, as it takes several minutes to transfer after each delivery and a new patient cannot be pulled up during the transfer. The large file size also renders the analysis extremely time consuming and nearly impossible to conduct for every treatment delivered. Thus, it would be ideal to analyze the compressed data. However, because of the compression, the compressed detector data do not match the instantaneous MLC shapes but their average over a projection. In order to analyze this compressed data, a new technique had to be developed. This technique uses an auto-associative kernel regression (AAKR) model. This technique uses the exit detector data from several historical treatment sinograms to develop the model. A new sinogram detector dataset is then applied to the model and the "true" or "error-free" value for each projection in this set is predicted. These predictions are compared to the values of the actual projections to determine if errors are present. Although the sinogram detector datasets have different attenuations, the model is capable of learning the relationships and accounting for these differences. Thus, when comparing the predicted and actual values, the only errors to appear result from errors in the delivery, such as errors in the MLC leaves or machine output.

4.2.1 AUTO-ASSOCIATIVE KERNEL REGRESSION (AAKR) THEORY

AAKR modeling is a novel nonparametric modeling technique developed by researchers at the University of Tennessee [Hines and Garvey 2006]. To understand auto-associative kernel regression, the basic principles of kernel regression must be understood. As discussed in the previous chapters, kernel regression (KR) is the process of estimating a parameter's value by calculating a weighted average of historical, exemplar observations. It is a well-known technique in statistics and empirical modeling. Generally, KR may be most compactly

represented by the so-called Nadaraya-Watson estimator. For a simple single-input, single-output (SISO) regression model, where the input *x* is used to estimate the output *y*, the Nadaraya-Watson estimator is as follows:

$$\hat{y}(x) = \frac{\sum_{i=1}^{n} K(X_i - x) Y_i}{\sum_{i=1}^{n} K(X_i - x)}, \qquad (4\text{-}1)$$

where: n is the number of exemplar observations in the KR model

X_i and Y_i are the input and output for the i^{th} exemplar observation

x is a query input

$K(X_i - x)$ is a weighting or kernel function, which generates a weight (similarity) for a given difference of a query from an exemplar vector

$\hat{y}(x)$ is an estimate of y, given x.

To fully explain KR, Equation 4-1 must be dissected. Overall, the equation performs 3 steps to arrive at its final prediction. For a query observation of the inputs, the KR estimation process first calculates the distance of the query from each of the input exemplars. An input exemplar is simply a historical observation of the variable being predicted. Thus, in this study, an exemplar is a fluence measurement found in the exit detector data used to construct a model. A query is then a fluence measurement from the exit detector data that is tested, or whose values are predicted by the model. The distance is simply a mathematical measurement of how far the query or the predicted fluence value is from the exemplar (the fluence values stored in memory). Next, the distances are supplied as inputs to a kernel function, which converts the distances to

weights (similarities). Finally, the weights are used to estimate the output by calculating a weighted average of the output exemplars. These steps are depicted in Figure 13.

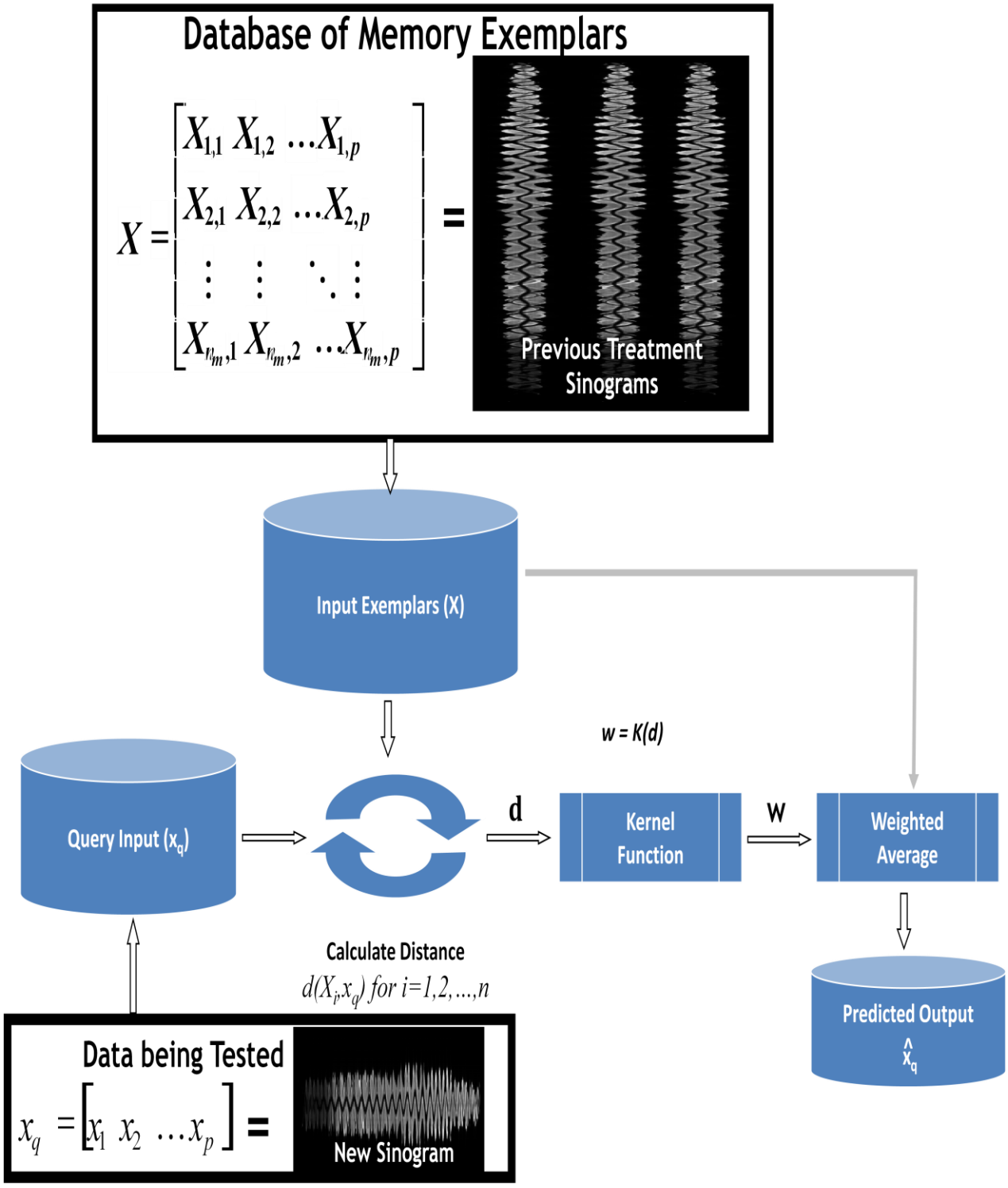


Figure 13 Diagram depicting the kernel regression process.

To understand the process shown in Figure 13, first consider the exemplar inputs **X** and outputs **Y**. These observations represent the "memory" of the KR model and are therefore often referred to as memory observations or memory vectors. In terms of the sinogram detector data, the memory vectors are the historical detector datasets used to develop the model. The input exemplars contain the sinogram detector data from past treatments. The columns in the input matrices represent projections in the historical detector data, while the rows represent each individual detector. Each observation or exemplar in **X** is then a fluence measurement (or detector count) from the particular detector for the given projection. In other words, each projection in the historical measured sinogram detector data is considered a variable, while fluence measurement in that projection is an observation.

Next, consider the query input $\mathbf{x_q}$. This query input $\mathbf{x_q}$ is a vector or matrix of the data being applied to the model. For this study, $\mathbf{x_q}$ contains the detector measurements for a given projection from the current treatment being analyzed. The observations in $\mathbf{x_q}$ are the KR model inputs and will be used to predict the output. To predict the output, it must first be determined how far the query is from the exemplar inputs. This is accomplished by evaluating a distance measure with each exemplar input and the query input as arguments. The distance calculation for the i^{th} exemplar observation is represented by $d(\mathbf{X}_i, \mathbf{x}_q)$. Notice that this calculation is repeated for each exemplar input, suggesting that the result of this entire operation is a vector of n distances **d**, which contains the distance of the query to all of the input exemplars.

At this point, the distances of the query to each of the input exemplars have been obtained and must be converted to weights or similarities. This is accomplished by evaluating a similarity or kernel function for each of the distances. For the vector of *n* distances **d**, the kernel function results in a vector of *n* weights **w**, which represent the similarity of the query to each of the input exemplars.

In the final step of the prediction process, the similarities of the query to each of the input exemplars are combined with the output exemplars to obtain estimates of the output. For KR, this is accomplished by calculating a weighted average of the output exemplars using the similarities of the query to the input exemplars as weighting parameters.

The KR prediction process can also be thought of as answering the following question: "Based on observed inputs **X** and outputs **Y**, what will be the system output for a new query input $\mathbf{x_q}$?" To answer this question, the following 2 questions must be answered first:

1) How similar is the query to the known inputs **X**?
2) The similarities of the query to the known inputs **X** can be used to measure which model outputs are most likely to occur. How do we aggregate these similarities to estimate the model output?

Now that the general process used in KR has been presented, Equation 4-1 can be revisited to see how each step is carried out, beginning with the distance calculation. The distance measurement used in Equation 4-2 can be seen as simply the difference of the input exemplar and the query:

$$d(X_i, x_q) = X_i - x_q. \qquad (4\text{-}2)$$

The specific distance function used in this research is the Euclidean distance**Error! Reference ource not found.**, which is also known as the L^2-norm. For a single input, the Euclidean distance for the i^{th} input exemplar and a query is given by the equation:

$$d(X_i, x_q) = \sqrt{(X_i - x_q)^2}. \qquad (4\text{-}3)$$

For p inputs, the Euclidean distance becomes:

$$d(\mathbf{X}_i, \mathbf{x}_q) = \sqrt{(X_{i,1} - x_{q,1})^2 + (X_{i,2} - x_{q,2})^2 + \cdots + (X_{i,p} - x_{q,p})^2} \qquad (4\text{-}4)$$

where: $\mathbf{X}_i$ is the i^{th} exemplar observation of the p inputs, and

$\mathbf{x_q}$ is the query observation of the p input variables, and x_{qj} is the query observation of the i^{th} input variable.

The distances calculated in Equation 4-4 are used to estimate the similarities via the kernel function K. These similarities can be interpreted as being weights and using the notation $\boldsymbol{w}$ to represent a vector of n weights as described above, the equation for the i^{th} weight (similarity) can be written as follows:

$$w_i = K(d(x_i, x_q)). \quad (4\text{-}5)$$

There are many kernel functions that are used to transform distances into similarities. Figure 12 depicts many of these functions. In general, a kernel function should have large values for small distances and small values for large distances. In other words, when a query point is nearly identical to a reference point, its distance should be small and, therefore, that particular reference point should receive a large weight and vice versa. Although each function shown in Figure 14 may have advantages or disadvantages in certain situations, Scott [1992] and Cleveland and Loader [1994] demonstrated in their studies that kernel function selection plays a non-critical role in the performance of locally weighted models.

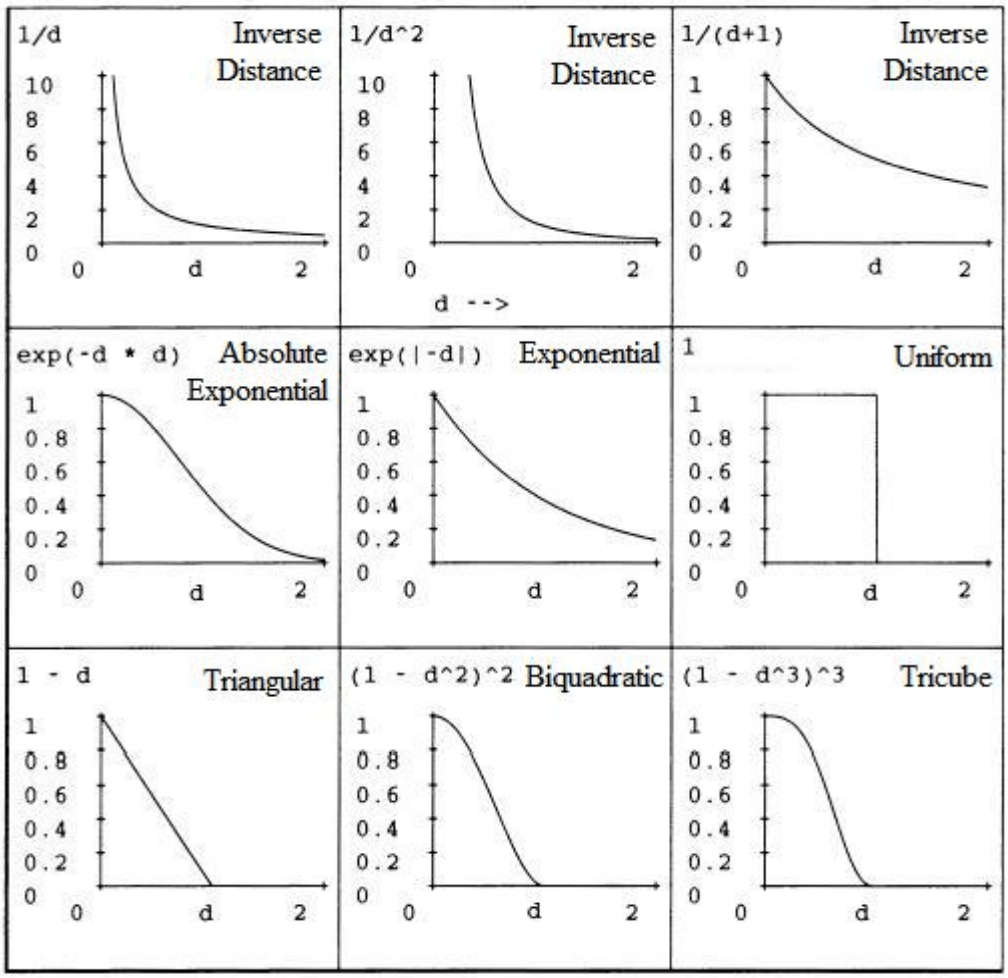


Figure 14 Examples of some common kernel functions.

The Gaussian kernel function (Figure 15) is an adequate selection, and was selected as the function for this research.

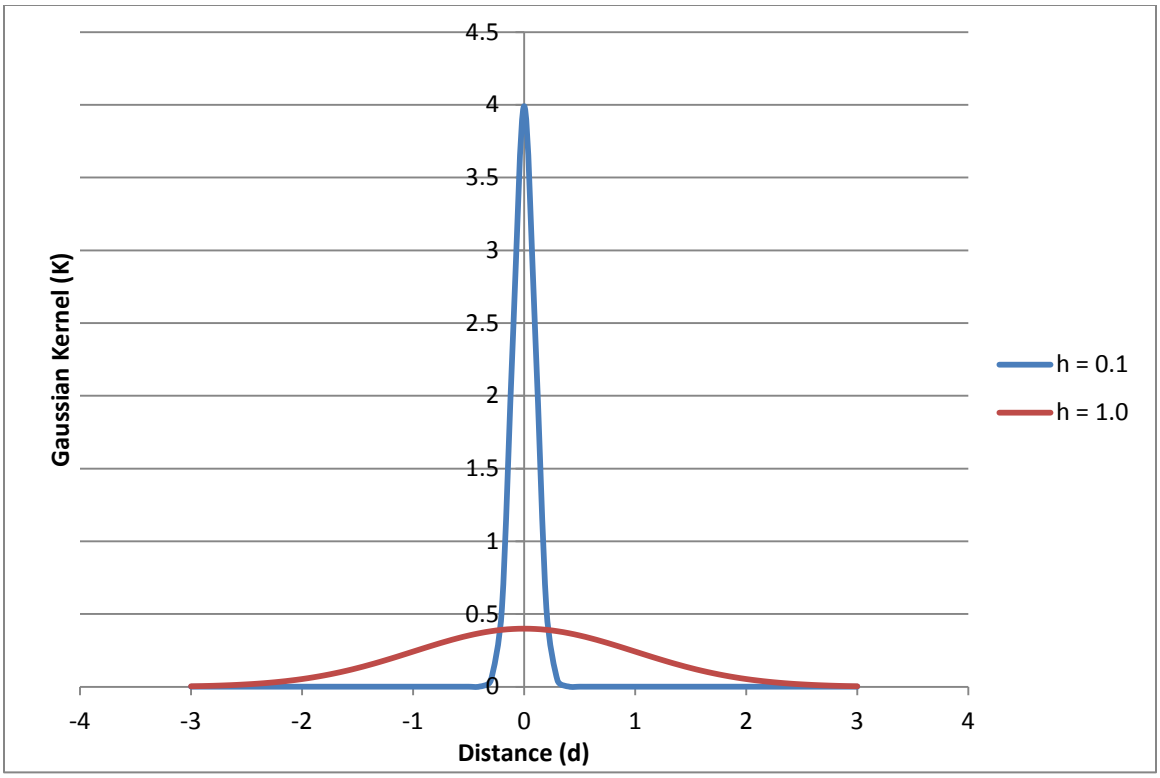


Figure 15 Example of the Gaussian kernel.

The Gaussian kernel is shown in the following equation:

$$K_h(d) = \frac{1}{\sqrt{2\pi h^2}} e^{-d^2/2h^2}. \qquad (4\text{-}6)$$

Here, h is commonly referred to as the kernel's bandwidth and used to control what effective distances are deemed similar. For mean-centered, unit-variance scaled data, the bandwidth generally has values on $0 < h \leq 1$. Figure 15 illustrates that the Gaussian kernel with the smaller bandwidth (h = 0.1) will only generate large weights when the distance is very close to 0, while the kernel with the larger bandwidth is less specific and will assign greater weights for a larger range of distances. For this application, a very small bandwidth is used, only a relatively small number of memory vectors (or fluence measurements from previous treatments stored in memory) will influence the estimated value of the fluence measurement from the treatment being

tested. However, if a very large bandwidth is used, there will be more significant weights, and a relatively large number of memory vectors will influence the estimated fluence value. Such a kernel function would then smoothen the AAKR's input-output relationship. Small bandwidths generally produce rough or non-smooth input-output relationships, while large bandwidths generally produce smooth input-output relationships. If the bandwidth is very small, there is a risk that noise in the data will be modeled. Conversely, a very large bandwidth will result in, overly smooth predictions, which renders detecting small errors in the measured detector data impossible. Therefore, the kernel must be optimized to identify the ideal bandwidth value.

After the Gaussian kernel converts the distances to similarities, the estimated similarities are to perform a weighted average of the output exemplars. Notice that in Equation 4-1, the sum of the weighted output exemplars is divided by the sum of the weights. This operation is simply a normalization that allows for the prediction to be represented as a "mix" of the output exemplars where each exemplar can have an influence of 0–1 (0–100%).

The methods for mapping an input query to similarities have already been discussed, and the different model architectures can now be examined in greater detail. There are 3 different KR architectures—inferential, heteroassociative, and auto-associative—that characterize the number and type of inputs and outputs.

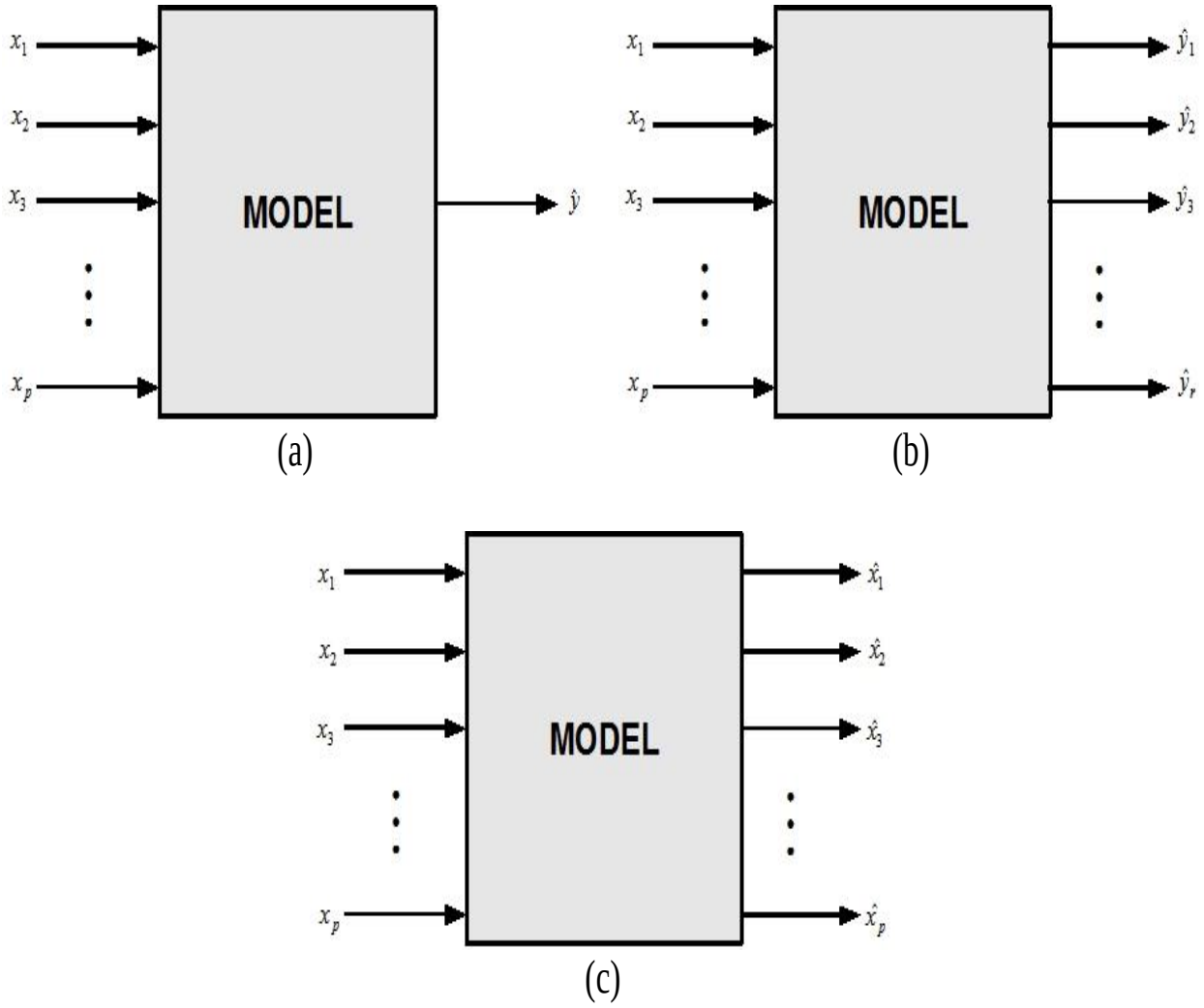


Figure 16 Illustration of (a) inferential, (b) heteroassociative, and (c) auto-associative model architectures.

As seen in Figure 16, an inferential model uses multiple inputs to infer an output, a heteroassociative model uses multiple inputs to predict multiple outputs, and an auto-associative model uses inputs to predict the "correct" values for the inputs, where "correct" refers to the relationships and behaviors contained in the exemplar observations. Thus, the auto-associative architecture is most suited for this application. For auto-associative kernel regression (AAKR), the inputs are "corrected" by performing a weighted average of the input exemplars:

$$\hat{\mathbf{x}}(x_q) = \frac{1}{a}\left[\sum_{i=1}^{n}(w_i X_{i,1}) \sum_{i=1}^{n}(w_i X_{i,2}) \cdots \sum_{i=1}^{n}(w_i X_{i,r})\right] \qquad (4\text{-}7)$$

where: the scalar a is defined as the sum of the weights, i.e., $a = \sum_{i=1}^{n_m} w_i$.

As shown in Equation 4-7, with an auto-associative architecture, a model's inputs and outputs are the same variables. These models have outputs trained to emulate the inputs over an appropriate dynamic range. Since data is often corrupted by noise, faults, and/or biases, an auto-associative model may be used to estimate the input's "correct" value. Here, "correct" refers to the values used to train the model. In other words, an auto-associative model will act to correct its input to those values that are in agreement with the correlations (relationships) embedded in the model during its training. A model's output will be a corrected version of the model input. If the inputs are faulty, this model should predict the un-faulted values under normal operating conditions. An AAKR model should be able to learn the relationships present between detector data from different treatments. Thus, when a model is developed with detector data from several treatments and tested with the current treatment, the correlations between the data should be normal unless there is some type of fault, such as an MLC error.

Auto-associative modeling is a new and innovative technique with wide application in the field of sensor fault detection. The AAKR model used in this research was developed and tested using code written in MATLAB version R2006b.

The memory matrix (X) and the kernel bandwidth (h) need to be optimized in an AAKR model. The developer must decide how many vectors to include in the memory matrix and how large to

make the bandwidth, which indirectly controls how many memory vectors are weighted heavily during prediction. Since the model's objective is to detect and quantify errors in the measured sinogram detector data, the model should be made as immune as possible to faulted data, which can be accomplished through optimization. Previous studies have investigated and reported optimization techniques [Hines and Usynin 2005; Wegerich 2002].

In this study, the AAKR model was optimized by performing a combinatorial grid search of the input parameter values. To begin the grid search, the user first defines the acceptable values for one or more of the model's initialization settings. Next, each combination of the initialization settings is used to construct a prototype model that is tested and scored. For this study, acceptable values for the kernel bandwidth were defined as 0.1–3 in increments of 0.1. This range is representative of typical kernel bandwidths; further, a bandwidth of 0.1 appeared to overfit the solution while a bandwidth of 3 over-smoothed it. Acceptable values for the number of memory vectors were defined as ranging from 1–543 in increments of 1, as there are a total of 543 fluence measurements in each projection. The grid search returned the model with the parameter combination that minimized the mean sum of the squared error (SSE) between the model predictions and the test data. These parameters were used in all future models. For this study, all of the vectors were included in the memory matrix and a bandwidth of 0.41 was used.

4.2.2 AAKR UNCERTAINTY

Since this technique could eventually be used to find errors in actual patient treatment, each model prediction must be accompanied by an estimate of its uncertainty. The inaccuracies in the training data and inherent limitations of the model contribute to the total uncertainty of these

techniques. More specifically, the selected predictor variables, the amount and selection of data used to develop the model, the model structure including its complexity, and the noise in the predictors and response (such as electrical contamination of the signal and board-induced, EMI (electromagnetic interference)- and RFI (radio frequency interference)-induced noise associated with the data acquisition) all influence the uncertainty. The uncertainty estimate reveals how accurate the model is and can also play a role in determining the threshold for error detection. The uncertainty of each AAKR prediction can be computed by using Monte Carlo methods [Efron 1993]0 or analytically determined [Gribok 2004]. For an approximate 95% confidence level, the analytical uncertainty for an AAKR prediction is computed with the following equation:

$$\hat{x}_q \pm 2\sqrt{\sigma_\varepsilon^2 + h^T h \sigma_\varepsilon^2 + Bias^2} \qquad (4\text{-}8)$$

where: σ_ε^2 is the variance of the error or noise in the measured sinogram detector data

h is the hat matrix of the weights given by $h = \left[\frac{w_1}{a} + \frac{w_2}{a} \cdots \frac{w_n}{a} \right]$ with $a = \sum_{i=1}^{n} w_i$

and the bias is defined as the model's systematic error and estimated by calculating the difference between the model's expected prediction, $E(\hat{x}_q)$, and an estimate of its "true" value $\mathbf{x_q}$.

Equation 4-8 gives the 95% prediction interval for an AAKR estimate. This prediction interval gives the upper and lower bounds between which there is a $(100 - \alpha) \times 100\%$ probability that the true output of the model prediction (at input x_q) lies. Unlike the commonly used confidence interval, the prediction interval is concerned with the confidence in the prediction of the targets.

The confidence interval only accounts for the variance component of the uncertainty and neglects the noise portion. Thus, the prediction interval is a much more practical parameter than the confidence interval because it provides the accuracy with which the desired response can be predicted, and not just the accuracy of the model itself.

4.2.3 SINOGRAM TESTING

When applying this AAKR model to the measured sinogram detector dataset, each of the predicted output vectors will be an error-corrected version of the input vectors. A properly developed model will produce predictions that show virtually no change when the corresponding input has been distorted by noise, faulty data, or missing data. This characteristic allows the auto-associative model to detect MLC errors or machine output errors by comparing each model input projection to the corresponding model estimates of the projections. If a projection that is entered into the auto-associative model is faulty due to an MLC positional error or machine failure, the model still produces a reasonably valid estimate of the correct projection as it obtains information from other correlated projections.

To test the ability of AAKR to detect errors in the measured sinogram detector data, the helical tomotherapy MLC controller files were extracted for 3 patients with tumors at different locations (1 head and neck, 1 prostate, and 1 lung) who had completed radiation treatment. The MLC controller files were modified by reducing the opening time of individual MLC leaves in specific projections by known amounts. Table 2 lists the inserted errors and their location in the delivery sequence for each case. The resulting error did not only depend on the percent reduction in the closing time of individual MLC leaves but also on how long the leaf was originally open in the

projection and its position in the given projection. The modified treatments are referred to as the "treatments with errors," and the unmodified treatments are referred to as "error-free." Treatment procedures with and without the known errors were manually created and delivered using the tomotherapy Calibration Data Interface. These treatments were delivered multiple times with different objects in the path of the beam in order to test the AAKR model's error detection capabilities in the presence of attenuation.

Table 2 Key for the errors that were intentionally inserted into the delivery sequence from actual patient treatment plans

	Head and Neck			Prostate			Lung		
% Error	MLC	Projection	Error (msec)	MLC	Projection	Error (msec)	MLC	Projection	Error (msec)
30	30	120	48.0	24	140	41.0	30	120	82.0
28	31	140	50.4	25	180	39.2	25	140	31.7
26	32	160	44.2	26	220	54.6	45	160	29.5
24	33	180	38.4	36	260	11.2	20	180	36.8
22	34	200	29.3	30	300	40.3	40	200	33.7
20	30	220	29.3	24	340	20.7	30	220	58.0
18	31	240	25.2	40	380	18.6	20	240	27.0
16	32	260	22.4	26	420	18.7	45	260	27.2
14	33	280	18.2	36	460	10.7	20	280	22.9
12	34	300	14.0	37	500	12.8	35	300	20.8
10	30	320	13.3	24	540	9.7	40	320	7.3
8	31	340	11.5	25	580	4.0	20	340	11.5
6	32	360	9.2	26	620	6.2	45	360	7.8

For both the prostate and head and neck treatments, a total of 6 delivery sequences were delivered. (1) No error with only the couch in the path of the beam, (2) error with only the couch in the path of the beam, (3) no error with a cylindrical phantom in the path of the beam, (4) Error with a cylindrical phantom in the path of the beam, (5) no error with an anthropomorphic phantom in the path of the beam, and (6) Error with an anthropomorphic phantom in the path of the beam. For the lung treatment, the same 6 delivery sequences were delivered, and additional

error and error-free treatments were delivered with a standard imaging (Middleton, WI) respiratory gating phantom in the path of the beam. The error and error-free lung plans were delivered with the phantom's longitudinal motion set to 5, 20, and 40 mm, with cycles ranging from 2–6 seconds. These combinations of amplitudes and phases were designed to replicate a wide range of different breathing patterns.

4.2.4 DOSIMETRIC IMPACT

The dosimetric impact of the MLC errors listed in Table 2 was evaluated for each of the test cases used in this study. The original unmodified dose distributions for each of the 3 test cases were extracted from an archived copy of the patients' plan. In addition, the files that define the MLC positions for dose calculation were also extracted. These files were then modified by inserting the known MLC errors for each test case. The doses were then recalculated with the modified MLC files using the tomotherapy planned adaptive dose reconstruction tool. Megavoltage CT (MVCT) images were not used during the dose reconstruction. The dose recalculation was performed using the same treatment planning CT dataset and contours as the original treatment plan. The only difference between the original treatment plan and the recalculated doses was the change in the MLC positions. Once the doses were recalculated, dose volume histograms (DVHs), isodose distributions, and dose difference maps were created using the planned adaptive tools.

4.3 RESULTS AND DISCUSSION

The auto-associative kernel regression (AAKR) modeling process described above is a new and "cutting edge" technology that has only recently been applied in industry. To date, AAKR or any similar modeling paradigm have never been used in the field of radiation therapy. Since AAKR is novel to radiation therapy, a straightforward study was conducted to test if it was a feasible technique that could help improve quality assurance. Multiple AAKR models were developed and tested by the investigators by using combinations of the measured detector data sets. This step is analogous to measuring a fluence intensity map during IMRT patient QA. The model was then optimized using the combination grid search method described above. The optimization process took nearly 3 minutes. However, optimization only had to be done once. All future models (using the selected optimal bandwidth and number of memory vectors) took less than a minute to develop. Models were developed using the error-free or unmodified sinogram detector data from the delivered treatments, and then tested using the detector data that had been modified to contain the known errors. However, models were never developed with data which had the same attenuation (or the same object in the path of the beam) as the dataset it was testing. This restriction verified that the model was robust to changes in attenuation, and could identify MLC and output errors even when the plan was delivered with different objects in the path of the beam. Figures 17, 18, and 19 present typical results for the AAKR models.

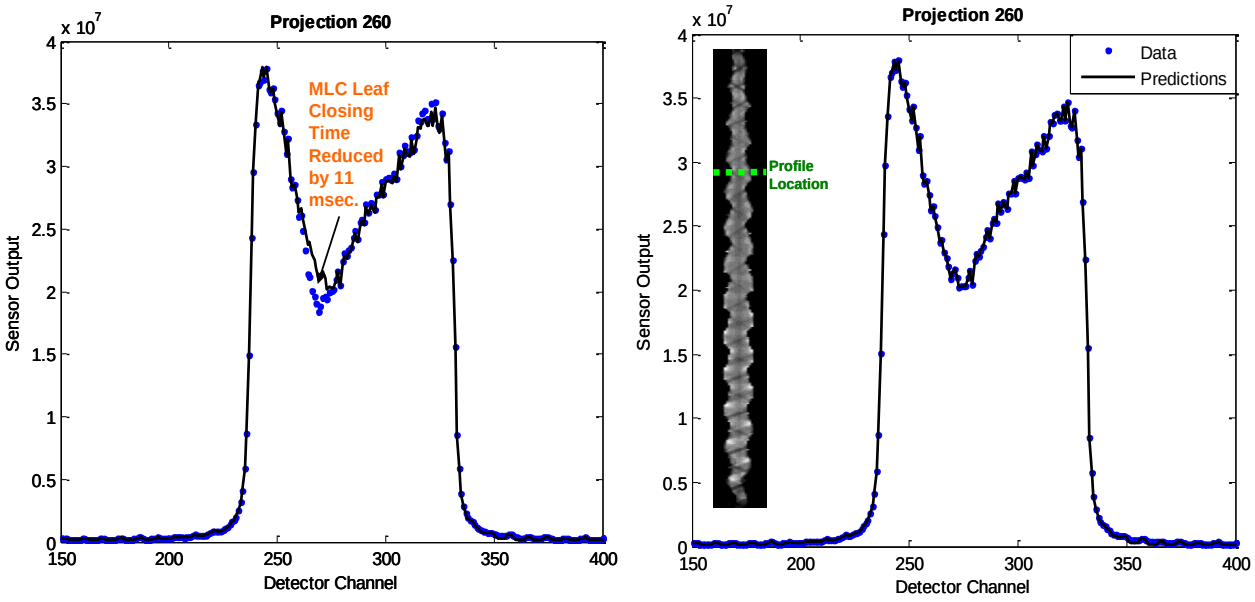


Figure 17 MLC positional error for the prostate test case (A) with and (B) without known errors.

As shown in Figure 17, the AAKR model correctly identified an error in the opening and closing of a leaf in projection number 260 for the prostate test case. In this example, the prostate treatment delivered with known errors had a single MLC leaf with an intentional decrease of 11 msec in the leaf opening time due to attenuation by an anthropomorphic phantom. This result proved that the AAKR technique correctly identified that a MLC leaf opened with a delay of 5.5 msec and closed 5.5 msec early from compressed exit detector data.

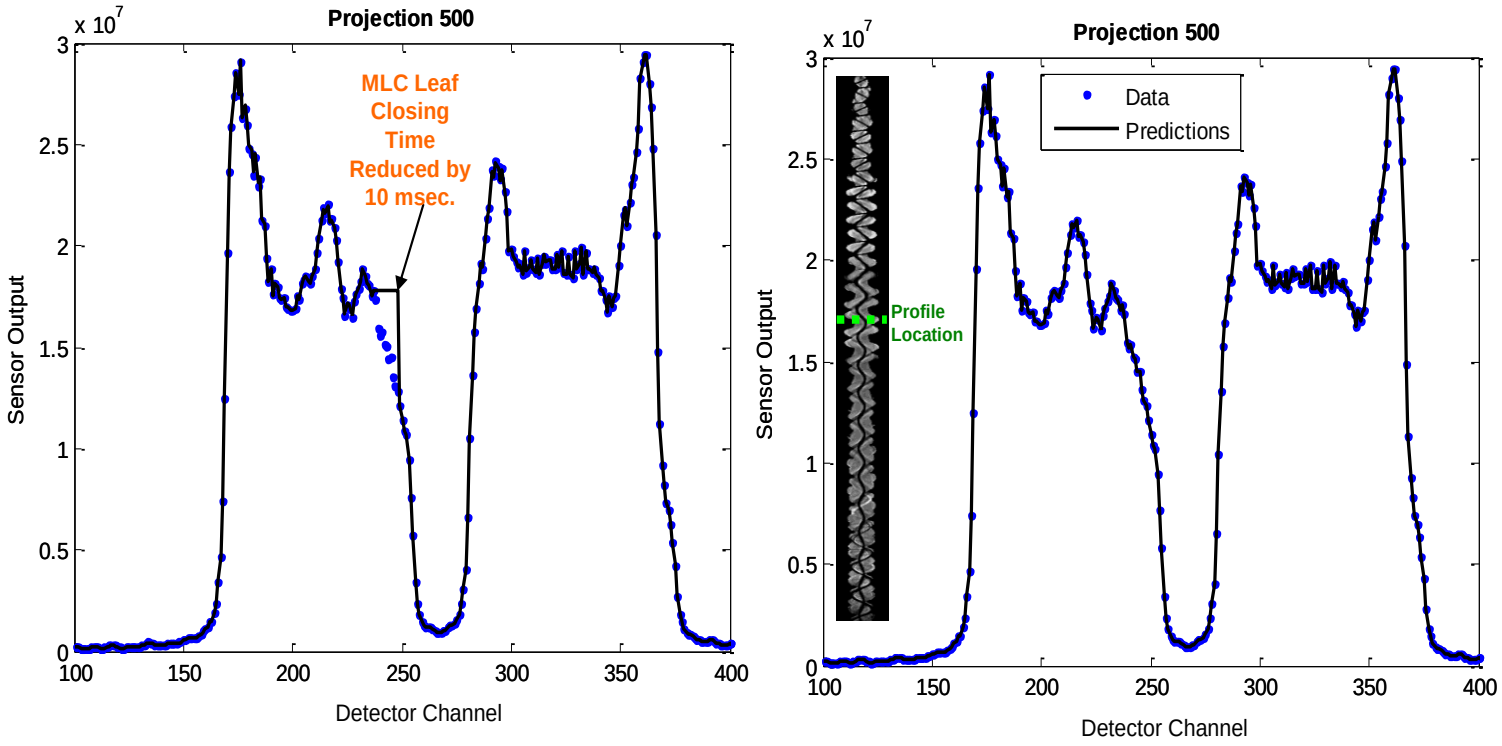


Figure 18 MLC positional error for the H&N test case (A) with and (B) without known errors.

Figure 18 shows similar results for the patient with the head and neck tumor. The model results shown in Figure 18 came from an AAKR model developed with sinogram detector data from the no error with only the couch in the path of the beam and no error with cylindrical phantom in the path of the beam deliveries. The results were then tested with the measured sinogram detector data from the error with an anthropomorphic phantom in the path of the beam delivery. In this example, the AAKR technique correctly identified MLC timing errors of 10 msec.

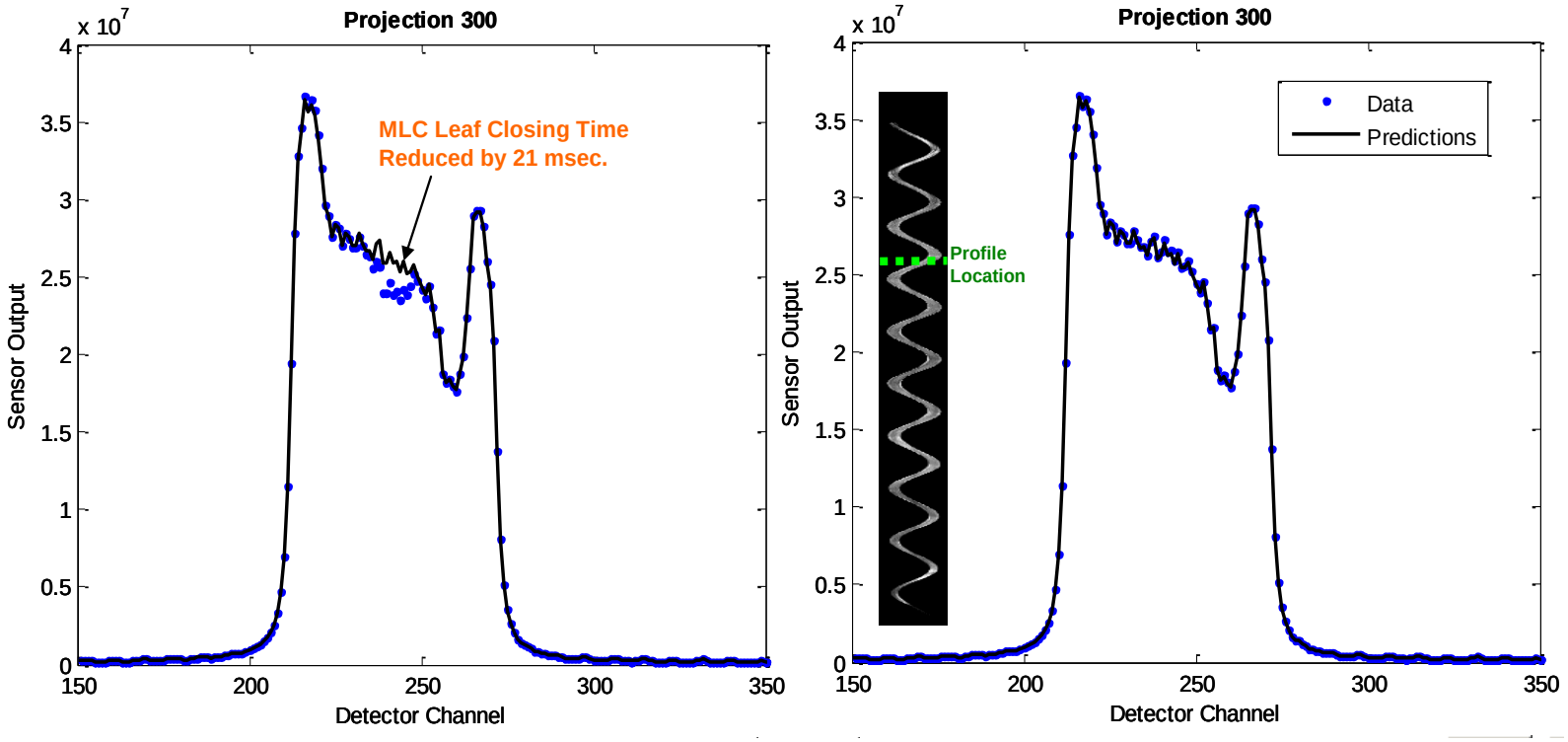


Figure 19 MLC positional error for the lung test case (A) with and (B) without known errors.

Figure 19 shows the result from a single projection of the error and error-free delivery to the lung. The model was developed with sinogram detector data from the no error with only the couch in the path of the beam and no error with a cylindrical phantom in the path of the beam deliveries. The model was then tested with detector data from the error and error-free deliveries with the motion phantom in the path of the beam. In this example, the motion phantom had its longitudinal motion set to 40 mm and the cycle set at 2 seconds. Using measured detector data from deliveries with stationary objects in the path of the beam to develop the model, and testing it with data from deliveries with motion illustrated that the model is not degraded by changes in the breathing pattern or patient motion that occurred during the treatment. The AAKR algorithm was still able to detect the MLC errors in the presence of simulated respiration. These figures are

only select highlights from the many models that were developed and tested in this study. The results presented are characteristic for all the cases tested.

The results shown in figures 17, 18, and 19 illustrate that the model correctly distinguished the MLC positional error from changes in attenuation caused by the phantoms. Furthermore, the model was capable of identifying errors in compressed detector data that had been summed over 70–100 frames. In general, errors greater than 7 ms were visually discernible by examining the model's prediction of the fluence values in comparison to the actual measured fluence values. Smaller errors could be detected, but the errors that were detected generally depended on the position of the erroneous leaf in the projection and the actual shape of the projection. In this study, only visual inspection of the predictions was used to identify errors. It is possible that more sophisticated statistical logic tests could be able to detect smaller errors in the prediction and also enable complete automation of this method. Determining the true threshold of error detection and applying statistical logic testing are potential areas for future research.

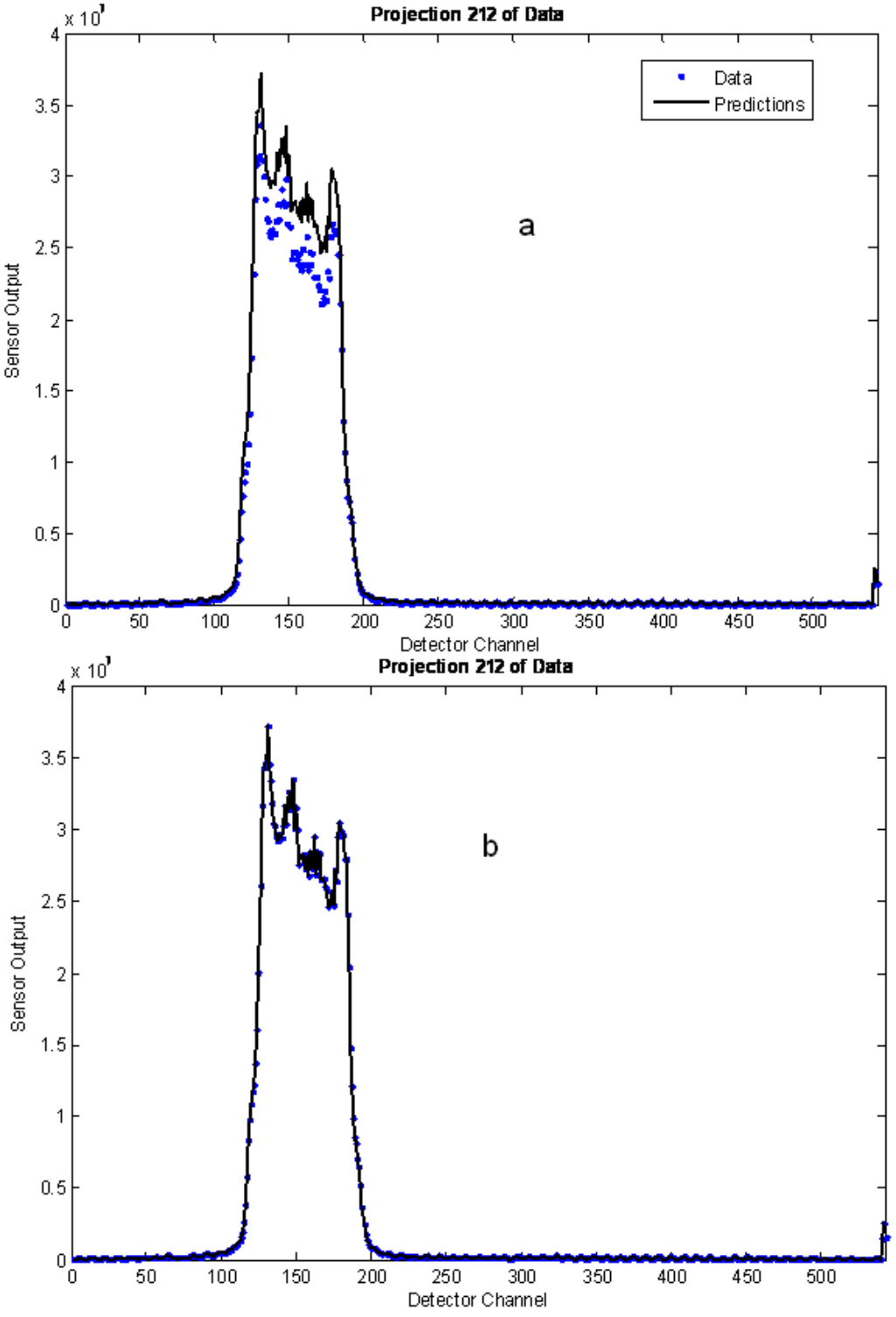


Figure 20 AAKR Model Predictions for Projection 212 of the prostate (a) Error with a cylindrical phantom in the path of the beam sinogram detector data and (b) No error with a cylindrical phantom in the path of the beam sinogram detector data illustrating how the model can detect machine output errors.

Figure 20 shows a different projection of a model testing detector data from the prostate "error with a cylindrical phantom in the path of the beam" delivery. In this figure, an output error occurred while the "no error with a cylindrical phantom in the path of the beam" treatment sequence was being delivered. This error was not intentionally inserted into the delivery sequence, and represents a real error that was detected by the model. The measured fluence was 13% lower than expected for all the MLC leaves in the high fluence region. The output of the system is measured and recorded using the ionization chambers located in the head of the accelerator. This data is stored in the detector data and was used to verify that there was a temporary decrease in machine output. The tomotherapy system will be interrupted if the output is 50% lower than the intended value for a duration of 1 second, or 5% lower than the intended value for 10 seconds. In this case, the output dropped by 13% for one projection (0.3 seconds), which would not trigger an interruption in the system. This type of error can be easily distinguished from an MLC error. In an individual MLC leaf error, only a few of the detectors (corresponding to the presence of the actual MLC leaf) would show mismatched values. In contrast, in the case of an output error, all the detectors in that projection deviate from their predicted values. The detection of this drop in output demonstrates that the AAKR model can detect real delivery errors.

The results shown in figures 17, 18, 19, and 20 confirm that when a projection input into the auto-associative model is faulty due to an MLC positional error or machine failure, the model still gives a reasonably valid estimate of the correct projection due to its use of information from other correlated projections. The results indicate that the AAKR model can learn complex detector data relationships, even with detector data with a low temporal resolution and beam

attenuation from the patient. However, in these figures, the model uncertainty is not shown. If the model uncertainty is large, there can be no confidence that the results seen in the figures are reflective of actual errors during the delivery. Rather, with large uncertainties, it becomes difficult to prove that a disagreement between a predicted value and the actual value is statistically significant.

To determine model uncertainty, the model must be tested using un-faulted data. Faulted data will inflate the uncertainty of an auto-associative model since the faulted test data is no longer representative of the training data. In fact, atypically large uncertainty values are indicative of errors and can be used to check for faults. In most auto-associative modeling schemes, an average uncertainty is calculated using data that is known to be good. This average uncertainty is then used as a baseline for comparison and to determine detectability.

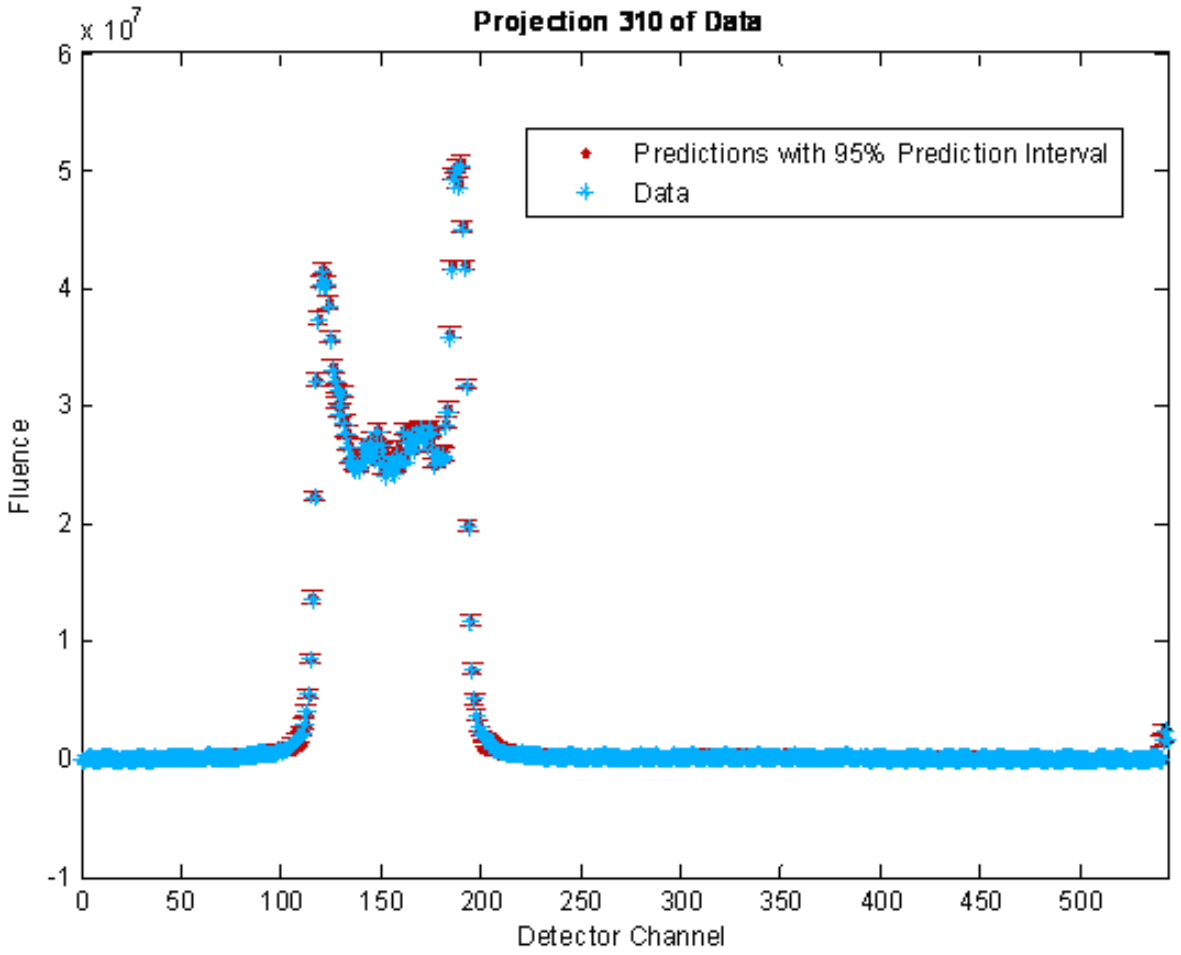


Figure 21 AAKR model predictions with their 95% prediction intervals for projection 310 of the prostate "Error with a cylindrical phantom in the path of the beam" sinogram.

Figure 21 shows the predicted fluence values for projection 310 of the prostate "error with a cylindrical phantom in the path of the beam" sinogram detector data accompanied by their 95% prediction interval. In this projection, no known error had been inserted into the delivery. Accordingly, no distinguishable errors were present in this sinogram, which is also reflected by the low uncertainty values. This low uncertainty is evidence of the model's accuracy. In the high fluence area (i.e., the area where the MLC leaves were open), the average uncertainty was 1.5% of the fluence value. This uncertainty value was found to be typical for the un-faulted projections, with the average uncertainty ranging from approximately 0.4–1.8%. The uncertainty

coverage for projection 310 was also calculated as 97.2%. Coverage refers to the percentage of the test data which fall within or are "covered" by the prediction interval (prediction ± uncertainty). For an accurate model testing un-faulted data, coverage is expected to be at least 95% to correspond to the 95% confidence level used in calculating the prediction interval. Coverage values exceeding 95% are evidence that the AAKR model is well-specified.

The low uncertainty values and good coverage values indicate that the AAKR model is extremely accurate in its predictions. They also suggest that this AAKR modeling scheme may be able to reveal delivery errors that cause the fluence to change by less than 2%. However, the model will only be able to detect delivery errors that are unique to the delivery being tested. Any errors that occur in a systematic manner at every treatment will be considered as normal operating conditions, and the system will therefore not be able to detect them.

The clinical impacts of the MLC errors measured in this study were tested using a dose re-calculation technique developed by the investigators. The original MLC sequence in the treatment planning system was modified to reflect the MLC errors. The dose was then recalculated for each case and compared against the error-free doses. In all 3 test cases, the re-calculated doses with the known MLC errors was considerably less than 1% (Figure 22). Therefore, it is possible for the AAKR technique to detect MLC errors before they reach clinical significance. One of the advantages of this technique is that it is independent of the tomotherapy MLC control system, which uses infrared tracking of the MLCs. This technique can be used to verify that the tomotherapy MLC interlock system is functioning correctly.

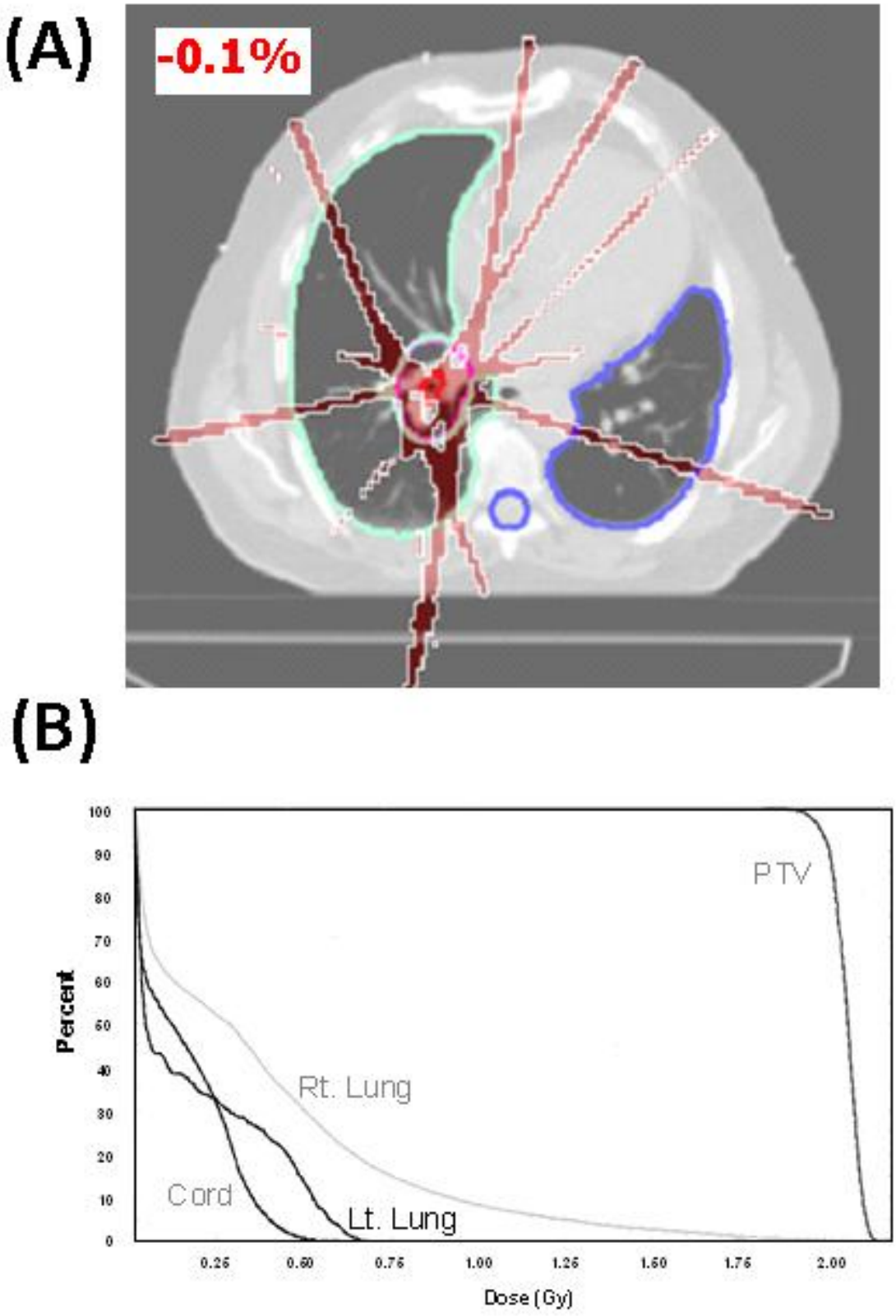


Figure 22 The 0.1% dose difference isodose lines and the DVH values for the lung case with errors versus the lung case without errors.

Overall, the results of this study suggest that AAKR modeling could be used to monitor and eventually improve the reliability of radiation delivery. In addition to analyzing sinogram detector data, the AAKR model could also be applied to EPID images acquired on other linear accelerator systems. For this application, the x and y pixels in the EPID images would be considered variables, while the actual detector counts would be considered observations. Using the same derivation, the model may be able to predict the "true" or "error-free" fluence values for each pixel.

However, it is important to reiterate that this AAKR modeling technique is only capable of detecting random errors. Any systematic errors will be masked and assumed to be the normal operating condition. Also, the technique cannot be applied until several fractions have been delivered. Thus, this technique cannot take the place of the standard IMRT patient specific QA. However, it is a valuable technique that provides valuable information about the integrity of the treatment. It still has the potential help the implementation of PART. It is possible a comparison between the planned vs. delivered sinogram could be conducted for the adapted plan in real time. Although this comparison will show some difference because of the variations in attenuation for the delivery, if the agreement was within 7% it could be safe to assume the adapted plan was valid. The AAKR technique could then later be used to detect finer errors in the adapted delivery.

4.4 SINOGRAM ERROR DETECTION CONCLUSIONS

A novel technique was developed for automatically evaluating exit dosimetry on a helical tomotherapy system using auto-associative nonparametric modeling. This technique was

capable of learning the complex detector data relationships, even in cases of detector data with low temporal resolution and beam attenuation from an anthropomorphic phantom. In the initial development of AAKR, the algorithm was successfully able to process highly compressed detector data and identify minute errors of single MLC leaves. Although the MLC errors that were intentionally introduced in this study have no dosimetric impact, being able to detect them illustrates that this method could provide early warnings for MLC leaf failures. AAKR was able to predict the correct or "error-free" values for a projection where a single MLC leaf decreased its opening time by 8 milliseconds. Moreover, the model also was able to determine machine output errors. The average uncertainty value for the un-faulted projections ranged from 0.4–1.8% of the detector signal. The low model uncertainty indicates that the AAKR model is extremely accurate in its predictions and also suggests that the model may be able to detect errors that cause the very slight changes in fluence. However, additional evaluation of AAKR is required to determine the minimum detectable error threshold from the compressed detector data. It is hoped that this technique will aid the implementation of PART.

5. CONCLUSIONS

Over the past decade, nonparametric empirical modeling techniques have gained popularity in numerous engineering-driven fields as viable prediction, smoothing, and state estimation tools. In general, the medical community, particularly the subsection involved in radiation therapy, has been slow to adopt these new techniques. However, these techniques are especially suited to this field, as their local characteristics are ideal for modeling patient data. This work has proven nonparametric modeling's usefulness in radiation therapy, as it relates to the development and implementation of Predictive Adaptive Radiation Therapy (PART).

The focus of this dissertation has been on introducing the concept of PART for lung tumors. To demonstrate PART, a nonparametric model was developed and tested on its ability to predict lung tumor response during image-guided radiation therapy treatment (IGRT). Volumetric computed tomographic (CT) images acquired by IGRT systems were used to measure tumor response over the treatment course. In this study, a patient database containing IGRT image data for 40 lesions obtained from patients who were imaged and treated with helical tomotherapy with doses ranging from 2.0–2.5 Gy per fraction was constructed. The changes in volume and mass over elapsed days were modeled using a LWR model. The LWR was able to predict future tumor volumes and masses as well as the associated confidence intervals based on observations from only the first two weeks of treatment. To test the predictive accuracy of the model, a leave-one-out cross-validation technique with the measured tumor responses was used. The findings showed that the absolute mean error for prediction of the mass and the final volume were 10.2% and 8.4% respectively. An addition of the patient data led to decrease in the absolute mean error. The greatest model uncertainty was recorded near the middle of the treatment, where the

predictors were varied and complex tumor response relationships. It can be concluded that since the predictions were accurate based on the quantified uncertainty, they can be used to adapt a patient's treatment and help decide if an intensity modulated integrated boost to the location of the residual tumor would be beneficial. Thus, PART is a feasible technique for lung radiotherapy.

However, adaptation of a treatment plan will require more patient specific QA, which could be a large drain on department resources. With that in mind, an auto-associative kernel regression (AAKR) model was applied to tomotherapy delivery sinograms with the goal of detecting errors during tomotherapy delivery. This particular modeling scheme is especially suited for monitoring the exit detector data. This model is capable of learning the relationships in the detector data, even when different degrees of beam attenuations and low temporal resolution as a result of patient's movement are recorded. Thus, the model should be able to distinguish random errors in the delivery (such as dropped pulses, or a misfiring MLC leaf) from the normal relationships caused by beam passing through the patient as he breathes. To determine the model's performance, multiple AAKR models were developed and tested using combinations of the measured detector data sets. However, models were never developed with data which had the same attenuation (or the same object in the path of the beam) as the dataset it was testing. This restriction verified that the model was robust to changes in attenuation. The developed models proved to be reliable, as they could predict the correct error free values especially for an auto-associative kernel regression (AAKR) model. Models of a projection that had "no known error" inserted into delivery, showed no distinguishable errors in the sonograms, which is evidence of the model's accuracy. On the other hand, comparison of the model predictions and measured

detector data for the modified projections (which was a projection where an MLC error had purposefully been inserted) made it clear that the model was correctly predicting the unfaulted fluence values. The low uncertainty values were further indicators of the model's accuracy. Thus, the AAKR model could serve as a prognositic tool for the tomotherapy system and provide warning of machine faults. However, it must be reiterated that this AAKR technique can never be the only form of machine QA because it will always consider errors that are occurring in a systematic manner as normal operating conditions and allow them to go undetected. But the technique shows promise as a prognostic tool. And it is hoped that when coupled with an on-line sinogram comparison method, it could still be used to ease the implementation of PART.

5.1 RECOMMENDATIONS FOR FUTURE WORK

There are several future directions for related work. In the lung response study, a promising area for further research is trying to correlate the predicted tumor response during treatment to clinical outcomes. However, for the correlation to be significant, a much larger database of patients is required, as well as many years of follow up data. It would also be interesting to add PET images acquired before, during, and after treatment into the model and see if this functional information will improve the model performance. Another area of future work is to extend PART beyond just lung tumors and investigate modeling different anatomical sites. However, only sites which have tumors whose response is visible on the daily IGRT images can be considered.

More research is also needed before the AAKR technique can be implemented clinically. Full automation of the technique is still an area warranting further research. It may even be possible

to employ the method in an on-board treatment computer. Once implemented, deciding how much of a reduction in traditional IMRT QA is appropriate and will not affect patient safety is another task that needs to be performed.

REFERENCES

Atkeson C, Moore A, and Schaal S (1996). Locally weighted learning. *Artificial Intelligence* 1996;11:11-73.

Barnes EA, Murray BR, Robinson DM,(2001). Dosimetric evaluation of lung tumor immobilization using breath hold at deep inspiration. *Int J Radiat Oncol Biol Phys* 2001;50:1091–1098.

Bowling, J. (2012), "Leksell Gamma Knife Treatment Planning via Kernel Regression Data Mining Initialization and Genetic Algorithm Optimization," PhD dissertation, Nuclear Engineering Department, University of Tennessee, Knoxville, January 2012.

Bradley JD and Deasy JO (2008), Nonlinear Kernel-Based Approaches for Predicting Normal Tissue Toxicities, *Machine Learning and Applications,* ICMLA '08, 2008.

Brewster L., Mohan R., Mageras G., Burman C., Leibel, S. and Fuks, Z. (1995), Three dimensional conformal treatment planning with multileaf collimators, *Int. J. Rad. Onc. Biol. Phys.* 33, 1081–1089, 1995.

Cleveland, W. S.; Loader, C. Smoothing by local regression: Principles and methods, Technical Report 95.3, AT&T Bell Laboratories, Statistics Department, Murray Hill, NJ: 1994.

Dirkx , J . van Sörnsen de Koste , S . Senan (2004). A treatment planning study evaluating a simultaneous integrated boost technique for accelerated radiotherapy of stage III non-small cell lung cancer. *Lung Cancer* 45(1):57 – 65; 2004.

Donnelly A, Misstear B, Broderick B (2011). Application of nonparametric regression methods to study the relationship between NO2 concentrations and local wind direction and speed at background sites. *Sci Total Environ.* 15,409(6):1134-44; 2011.

Dougherty, G. (2009), *Digital Image Processing for Medical Applications*. Cambridge University Press; 1st edition, May 2009.

Efron, B.; Tibshirani, R.J. *An Introduction to the Bootstrap*, Chapman and Hall, New York, NY: 1993.

Ezzell, GA, Galvin, JM, Low, D (2003), Guidance document on delivery, treatment planning, and clinical implementation of IMRT: report of the IMRT Subcommittee of the AAPM Radiation Therapy Committee. *Med Phys.* 30(8):2089-2115, 2003.

Ezzell GA, Burmeister JW, Dogan N, LoSasso TJ, Mechalakos JG, Mihailidis D, Molineu A, Palta JR, Ramsey CR, Salter BJ, Shi J, Xia P, Yue NJ, and Xiao Y.(2009), "IMRT commissioning: Multiple Institution planning and dosimetry comparisons, a report from AAPM Task Group 119," *Med. Phys*. 36:5359–5373, 2009.

Fan J. and Gijbels I (1996). *Local Polynomial Modelling and its Applications*. Chapman and Hall, London: 1996.

Feng, M and A Eisbruch (2007), Future issues in highly conformal radiotherapy for head and neck cancer. *Journal of Clinical Oncology*. 25, 1009-1013, 2007.

Ferris M and Voelker M. Fractionation in radiation treatment planning. *Math. Program.*, Ser. B 101:387–413, 2004.

Fox J (2010). *Multiple and Generalized Nonparametric Regression***.** *SAGE Publications,* 2010.

Galvin, J.M., Han, K., Cohen, R. (1998), A comparison of multileaf-collimator and alloy-block field shaping. *Int J Radiation Oncology Biol Phys.* 40, 721-731, 1998.

Garvey, D.R. (2006), "An Integrated Fuzzy Inference Based Monitoring, Diagnostic, and Prognostic System," PhD dissertation, Nuclear Engineering Department, University of Tennessee, Knoxville, December 2006.

Gerard PD and Schucany WR (1999). Local Bandwidth Selection for Kernel Estimation of Population Densities with Line Transect Sampling. *Biometrics*, 1999; 55: 769-773.

Godley A, Ahunbay E, Peng C and Li XA (2009), Automated registration of large deformations for adaptive radiation therapy of prostate cancer. *Medical Physics* 2009; 36 1433-41.

Grills IS, Yan D, Martinez AA, Vicini FA, Wong JW, Kestin LL (2003), Potential for reduced toxicity and dose escalation in the treatment of inoperable non-small-cell lung cancer: a comparison of intensity-modulated radiation therapy (IMRT), 3D conformal radiation, and elective nodal irradiation. *Int J Radiat Oncol Biol Phys* 57(3):875-890, 2003.

Gribok, A.V., A. M. Urmanov, and J.W. Hines. Uncertainty Analysis of Memory Based Sensor Validation Techniques. *Real Time System*. 2004 ; 27 : 7-26.

Haas, Olivier C. L. (2008), *Intelligent and Adaptive Systems in Medicine*. Taylor & Francis Group, 2008.

Hall, Eric (2006), *Radiobiology for the Radiobiologist*. Lippincot, 4th edition, 2006.

Harris, Carley Elizabeth, "An Automated Diagnostic Tool for Predicting Anatomical Response to Radiation Therapy." PhD diss., University of Tennessee, 2009.

Haupt, R. L. and Haupt, S. E. *(1998), Practical Genetic Algorithms.* John Wiley & Sons, NewYork: 1998.

Hendee, W., Ibbott, G., & Hendee,E. (2005). *Radiation Therapy Physics*. Wiley-Liss, 3rd edition, 2005.

Hines, J.W., L.W. Townsend, and T. F. Nichols (2005), "SPE Dose Prediction Using Locally Weighted Regression", *Radiation Protection Dosimetry*, 2005; 114, No. 1-4, pp 131-134.

Hines, J.W.; Usynin, A.A. (2005). MSET performance optimization through regularization. *Nuclear Engineering and Technology* 2005; 37:177–184.

Hines, J.W. and Garvey, D.R. (2006), Development and Application of Fault Detectability Performance Metrics for Instrument Calibration Verification and Anomaly Detection, Journal of Pattern Recognition Research, 1, 2-15, 2006

Hines, J.W.; Seibert, R.M.; Technical Review of On-line Monitoring Techniques for Performance Assessment: Part I State of the Art. NUREG/CR-6895, U.S. Nuclear Regulatory Commission, Washington, D.C., 2006.

Hines, J. W.; Seibert, R.M.; Garvey, D.D.; Usynin, A.A.; Arndt, S.A.; Technical Review of On-line Monitoring Techniques for Performance Assessment: Part II Theoretical Issues. NUREG/CR-6895. U.S. Nuclear Regulatory Commission, Washington, D.C. 2007.

Hinderer R. (2003). *Development of an efficient detector system for megavoltage photons*. PhD Thesis University of Wisconsin, Madison, 2003.

Holmes T., Mackie TR., (1994), A comparison of three inverse treatment planning algorithms. *Phys Med Biol*, 39(1), 91-106, 1994.

Humberstone, Matthew (2010), "An Adaptive Nonparametric Modeling Technique for Expanded Condition Monitoring of Processes. " ," PhD dissertation, Nuclear Engineering Department, University of Tennessee, Knoxville, December 2010.

Karzmark C.J. and Morton R.J. (1981), *A primer on theory and operation of linear accelerators in radiation therapy*. Bureau of Radiological Health, FDA 82-8181, December 1981

Keller W, Tome M, Ritter A, and Mackie TR. Design of adaptive treatment margins for non-negligible measurement uncertainty: application to ultrasound-guided prostate radiation therapy. *Phys. Med. Biol.*, 49:69–86, 2004.

Kim D, Shyr Y, Chen H, et al (2005). Response to combined modality therapy correlates with survival in locally advanced non–small-cell lung cancer *Int J Radiat Oncol Biol Phys* 2005; 63: 1029-1036.

Kissick MW, Fenwick J, James JA, (2005), The helical tomotherapy thread effect. *Medical Physics,* 32(5),1414–1423, 2005.

Khan, Faiz M. (2003), *The Physics of Radiation Therapy*. Lippincott Williams & Wilkins, 3rd edition, 2003.

Kron T, Grigorov G, Yu E, *et al* (2004). Planning evaluation of radiotherapy for complex lung cancer cases using helical tomotherapy. *Phys Med Biol* 2004;49:3675–3690.

Kupelian PA, Ramsey CR, Meeks SL, *et al* (2005). Serial megavoltage CT imaging during external beam radiotherapy for non–small-cell lung cancer: observations on tumor regression during treatment. *Int J Radiat Oncol Biol Phys* 2005; 63: 1024-1028.

Lai, T.L. and Xing, H. (2008). Statistical *Models and Methods for Financial Markets*. Springer-Verlag, New York 2008.

Langen KM, Papanikolaou N, Balog J, Crilly R, Followill D, Goddu SM, Grant W 3rd, Olivera G, Ramsey CR, Shi C (2010), QA for helical tomotherapy: report of the AAPM Task Group 148. *Medical Physics*. 2010 Sep;37(9):4817-53.

Li R., Xing L. (2011), A kernel method for real-time respiratory tumor motion estimation using external surrogates, *Proc. IEEE ICMLA*, Honolulu, HI, USA, 2011.

Louwe, R.J.; Tilenburg, R.; van Ingen, K.M.; Mijnheer, B.J.; van Herk, M.B. The stability of liquid-filled matrix ionization chamber electronic portal imaging devices for dosimetry purposes. *Med Phys* 2004; 31:819–827.

Mackie T R, Holmes T W, Swerdloff S, Reckwerdt P J, Deasy J O, Yang J, Paliwal B R and Kinsella T J. (1993). Tomotherapy: a new concept in the delivery of dynamic conformal radiotherapy. *Medical Physics*. 20, 1709–1519, 1993.

Mackie T R, Holmes T W, Reckwerdt P J ,and Yang J N (1995). Tomotherapy: optimized planning and delivery of radiotherapy. *International Journal of Imaging Science Technology.* 6, 43–55, 1993.

Mackie T R (2006). History of tomotherapy. *Phys. Med. Biol.* 51:427-449; 2006

Marks LB, Yorke ED, Jackson A, Ten Haken RK, Constine LS, Eisbruch A, Bentzen SM, Nam J, Deasy JO (2010). Use of normal tissue complication probability models in the clinic. *Int J Radiat Oncol Biol Phys*. 2010 Mar 1;76(3 Suppl):S10-9.

Martinez D, Yan D, Lockman D, Brabbins K, Kota M, Sharpe DA, Jaffray F, and Wong J. (2000). Improvement in dose escalation using the process of adaptive radiotherapy combined with three-dimensional conformal or intensitymodulated beams for prostate cancer. *IntRad*, 48:289–302, 2000.

Moran J., Dempsey M., Eisbruch A., Fraass B., Galvin B., Ibbott G., Marks L. (2011). Safety considerations for IMRT: Executive summary**.** *Practical Radiation Oncology 1,* 190-195, 2011.

Molineu A, Followill D.S., Balter, P.A., Hanson W.F., Gillin M.T., Huq M.S., Eisbruch, A, Ibbott G.S. (2005). Design and implementation of an anthropomorphic quality assurance phantom for intensity-modulated radiation therapy for the radiation therapy oncology group. *International Journal of Radiation Oncology Biology Physics* 63,577–583, 2005.

Palta, JR, Mackie, TR (2003), *Intensity-Modulated Radiation Therapy: The State of the Art.* Medical Physics Monograph No. 29. Published for the American Association of Physicists in Medicine 2003 Summer School Proceedings, 2003.

Press WH, Flannery BP, Teukolsky SA, and Vetterling WT (1992). *Numerical Recipes in FORTRAN: The Art of Scientific Computing*. Cambridge University Press, Cambridge, England: 1992.

Ramsey CR, Seibert RM, Mahan SL, Desai D, Chase D (2006), Out-of-field dosimetry measurements for a helical tomotherapy system. Journal of Applied Clinical Medical Physics 7(3), 1–11, 2006.

Rasmussen, B. and J.W. Hines (2005), On-Line Sensor Calibration Monitoring Uncertainty Estimation, *Nuclear Technology*, Sept, 2005, 151(3).

Sarokin DA and Schulkin JA (1994). Environmental Justice: Co-Evolution of Environmental Concerns and Social Justice. The Environmentalist 14, 1994.

Scrimger RA, Tome WA, Olivera GH, *et al.* Reduction in radiation dose to lung and other normal tissues using helical tomotherapy to treat lung cancer, in comparison to conventional field arrangements. *Am J Clin Oncol* 2003;26: 70–78.

Scott, D. W. *Multivariate Density Estimation*, Wiley, New York, NY: 1992.

Seibert RM, Ramsey CR (2007), Respiration Motion Prediction using Time-delay Kernel Regression Modeling. *International Journal of Radiation Oncology, Biology, Physics* 69 (3), 2007.

Suntharalingam N, Podgorsak EB, and Hendry JH, Basic Radiobiology. In: Podgorsak EB ed. (2005), *Radiation Oncology Physics: A Handbook for Teachers and Students*. Vienna: IAEA, 502, 2005.

Tamhane AC, and Dunlop DD, *Statistics and Data Analysis from Elementary to Intermediate*. Prentice Hall, Upper Saddle River, NJ: 2000.

Takahashi S (1965), Conformation radiotherapy. Rotation techniques as applied to radiography and radiotherapy of cancer. *Acta Radiol Diagn Suppl. 242*, 1965.

Underberg R, Lagerwaard F, Slotman B, et al. Benefit of Respiration-Gated Stereotactic Radiotherapy for Stage I Lung Cancer: An Analysis of 4DCT Datasets. *Int J Radiat Oncol Biol Phys* 2005;62:554–560.

Underberg WM, Lagerwaard FJ, Tinteren HV, *et al*. Time trends in target volumes for stage I non-small-cell lung cancer after stereotactic radiotherapy. *Int J Radiat Oncol Biol Phys* 2006; 64: 1221-1228.

Velten, K. (2009), *Mechanistic Models II: PDEs, in Mathematical Modeling and Simulation: Introduction for Scientists and Engineers*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim,2009.

Veldeman, T., Madani, I., F. and De Neve, W. (2008). Evidence behind use of intensity-modulated radiotherapy: a systematic review of comparative clinical studies. *Lancet Oncology*, 9(4), 367-375, 2008.

Warkentin, B.; Steciw, S.; Rathee, S.; Fallone, B.G. Dosimetric IMRT verification with a flat-panel EPID. *Med Phys* 2003; 30:3143–3155.

Wand MP and Jones MC. *Kernel Smoothing, Monographs on Statistics and Applied Probability*. Chapman & Hall, London: 1995.

Wegerich, S (2002). Performance comparison of variable selection and grouping algorithms. SmartSignal Corp. Technical Report for Funded Research Project: "MSET Variable Selection Optimization", August 2002.

Wolpert D.H., Macready W.G. (1997), No Free Lunch Theorems for Optimization, *IEEE Transactions on Evolutionary Computation*, (1) 67,1997.

Williams, Francis Henry (1902). Roentgen rays in medicine and surgery. Macmillan, 1902.

Wu C, Jeraj R, Olivera G.H. and Mackie T.R. (2002) Re-optimization in adaptive radiotherapy, Phys. Med. Biol. 47:3181-3195, 2002.

Xing L.,Thorndyke B., Schreibmann E., Li L. (2006). Overview of image guided radiation therapy (IGRT), *Medical Dosimetry* 31, 91-112, 2006.

Yan, D, Vicini, F, Wong, J, Martinez, A (1997), *Adaptive radiation therapy*. Phys. Med. Biol. 42:123-132, 1997.

APPENDIX A. PATIENT TABLES

Table A-I. Patient Information

Lesion #:	Age:	Dose:	Stage:	TNM:	Pathology:	Chemotherapy:	Current Survival	Survival at Time of Death
1	73	60	II	T1N0M0	Adeno	No	20	--
2	56	62	I	T1N0M0	Adeno	Concurrent	31	--
3	81	64	I	T1N0M0	Adeno	No	14	--
4	65	70	I	T1N0M0	Adeno	Concurrent	--	--
5	67	70	I	T1N0M0	Squam	No	--	--
6	59	60	IIA	T1N1M0	Squam	No	--	2
7	79	65	I	T1N0M0	Squam	No	--	--
8	59	60	IIA	T1N1M0	Squam	No	--	2
9	78	65	IB	T2N0M0	Squam	Yes	10	--
10	73	68	I	T1N0M0	Squam	No	--	--
11	81	50	I	T1N0M0	Small	No	--	1
12	89	63.5	I	T1N0M0	Adeno	No	40	--
13	78	65	IB	T2N0M0	Squam	Concurrent	10	--
14	77	63.5	IV	T2NXM1	Adeno	No	--	13
15	59	68	IIB	T1N1M0	Squam	Concurrent	--	--
16	70	64	IB	T2N0M0	Adeno	No	11	--
17	59	70	IV	T3N0M1	Squam	Induction	7	--
18	59	74	IIIB	T3N0M0	Squam	Concurrent	23	--
19	76	60	I	T1N0M0	Squam	No	-	76
20	73	60	IB	T2N0M0	Unknown	No	3	--
21	78	63.5	I	T1N0M0	Adeno	No	--	--
22	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	--	--
23	51	58	IIIB	T3N0M0	Unknown	None	---	5
24	85	60	IIIA	---	Large Cell	---	--	--
25	68	60	IIIB	---	Adeno	---	--	--
26	67	60	IIIB	---	Squam	---	---	---
27	76	60	IIIA	---	Squam	---	--	--
28	60	60	IIIA	---	Adeno	---	--	--
29	76	60	IIIB	---	Large cell	---	---	---
30	70	60	IIIB	---	Adeno	---	--	--
31	45	50	IV	---	Adeno	---	--	--
32	60	60	IV	---	Squam	---	---	---
33	76	63	IIIB	---	Squam	---	--	--
34	71	64	IIIA	---	Adeno	---	--	--
35	70	60	IIIA	---	Squam	---	---	---
36	70	60	IIIA	---	Large cell	---	--	--
37	71	60	IIIA	---	Large cell	---	--	--
38	77	60	IIIA	---	Large cell	---	---	---
39	77	50	IIIB	---	Adeno	---	--	--
40	76	62	IIIA	---	Squam	---	--	--

The Following are the same patients
Lesions 1 & 2
Lesions 7 & 9
Lesions 10 & 14

Table A-II. Predicted Volumes for Last Fraction

Lesion	Initial Volume (cc)	Actual Final Volume (cc)	Percent Change in Mass (%)	Percent Change in Volume (%)	Predicted Final Volume (cc)	Error between Predicted and Actual (cc)	± 95% Confidence Interval (cc)
1	1.90	1.80	-24%	-5%	1.9	0.1	1.0
2	2.10	0.90	-69%	-57%	0.8	0.1	0.2
3	3.20	1.90	-86%	-41%	1.9	0.0	0.3
4	4.80	1.00	--	-79%	1.2	0.2	0.6
5	5.70	3.70	--	-35%	4.0	0.3	0.8
6	5.90	3.90	-25%	-34%	4.2	0.3	0.5
7	8.10	1.90	--	-77%	1.4	0.5	0.2
8	8.90	2.40	-91%	-73%	2.2	0.2	0.3
9	13.20	2.80	-57%	-79%	3.4	0.6	0.6
10	13.70	8.90	--	-35%	9.8	0.9	1.0
11	18.30	8.20	-42%	-55%	7.7	0.5	0.6
12	19.10	7.80	-63%	-59%	8.6	0.8	1.3
13	21.30	11.60	-76%	-46%	12.1	0.5	3.3
14	21.70	11.00	-54%	-49%	11.8	0.8	0.9
15	21.90	7.30	--	-67%	7.7	0.4	0.6
16	22.60	7.30	--	-68%	8.3	1.0	1.0
17	23.50	15.90	-46%	-32%	15.7	0.2	2.3
18	24.70	2.80	-65%	-89%	2.2	0.6	1.1
19	32.60	14.30	-75%	-56%	15.2	0.9	0.6
20	33.50	13.10	-68%	-61%	13.5	0.4	0.8
21	33.80	12.40	-35%	-63%	11.4	1.0	1.5
22	35.73	27.46	--	-23%	28.9	1.5	1.7
23	37.55	31.74	--	-15%	32.1	0.4	1.6
24	41.20	23.90	--	-42%	25.2	1.3	1.6
25	50.50	31.00	--	-39%	31.9	0.9	0.6
26	65.04	57.26	--	-12%	53.5	3.8	8.0
27	65.90	20.00	--	-70%	21.1	1.1	2.0
28	99.90	33.50	--	-66%	31.8	1.7	1.7
29	104.70	75.20	--	-28%	70.0	5.2	6.0
30	108.43	64.02	--	-41%	64.9	0.9	2.7
31	115.30	89.70	--	-22%	83.8	5.9	6.8
32	132.30	90.30	--	-32%	92.1	1.8	3.1
33	155.80	74.60	--	-52%	76.0	1.4	2.3

Table A-II. Predicted Volumes for Last Fraction (continued)

Lesion	Initial Volume (cc)	Actual Final Volume (cc)	Percent Change in Mass (%)	Percent Change in Volume (%)	Predicted Final Volume (cc)	Error between Predicted and Actual (cc)	± 95% Confidence Interval (cc)
34	171.30	140.10	--	-18%	130.3	9.8	9.7
35	173.50	108.40	--	-38%	112.9	4.5	1.1
36	201.70	25.60	--	-87%	37.6	12.0	11.4
37	250.00	138.80	--	-44%	145.4	6.6	8.1
38	484.60	386.14	--	-20%	361.8	24.3	25.0
39	735.80	397.00	--	-46%	354.4	42.6	44.7
40	737.20	265.20	--	-64%	261.8	3.4	15.9

APPENDIX B. MATLAB CODE

```
function [rms ypred yact] = lungcom(x, h_KR)
l=[1:20]; %days used for observation (remember that l=1 goes w/ day=0)
for i=1:size(x,2) %iterate through patients using other patients to predict
for B=1:50 % start bootstrap for uncertainty
g = sort(randsample(l,size(l,2),1)); %get bootstrap samples
[row, col]=size(x(g,:));
yact=x(66,:); %row 66 has the final volume for each patient at the end of
their treatment
h_KR=3.5; %found h_KR using LVOCV
a=[1:col]; %used to set difference
V=diag(ones(20,1)); %for variance
b=setdiff(a,i); %set difference to get vector of all patients except the one
that is being predicted
[ys,ym,yst]=zscore(x); %scale the data
XT=ys(g,b); %set x-training data
XQ=ys(g,i); %set x-query data
YT=ys(66,b); %set y-testing data
YQ=ys(66,i); %set y-query data
p=1;
[YHAT1] = lwr_mm(XT',YT',h_KR,XQ'); %predict final volume
ypred(B,i)=unscore(YHAT1, ym(i), yst(i)); %unscale the predicting
end
end
rms=mean(abs(mean(ypred)-yact));
end

function [yq unc] = lwpr1_mm_local(X,Y,sc,q, noisevar)
%
%   Locally Weighted Polynomial Regression with Local Gaussian kernal and
Euclidean distance.
%   Multivariate input matrix.
%
% X     - input data matrix 19X10
% Y     - output data matrix
% sc    - bandwidth vector
% q     - matrix of query vectors: row = observation, column = inputs
% yq     - prediction for q

% Hines 2002
% University of Tennessee
X=[X X.^2]; % compute squares of inputs
q=[q q.^2];
%[X xm xs]=zscore1(X);
%q=zscore1(q,xm,xs);

[n,nx] = size(X);           % n is number of patterns
[n,ny] = size(Y);         % nx and ny are number of inputs
                            % and outputs


iters=size(q,1);
yq=zeros(1,iters);
```

```
%h = waitbar(0,'Please wait...');

Xq=X;                          % Original Inputs
X = [X ones(n,1)];             % Inputs appended with ones for intercept
nx = nx+1;                     % New number of inputs
scn=[sc sc.^2];
scn=[scn 1];
for i=1:iters;
     q1 = [q(i,:) 1]; % Query appended with ones
     B=repmat(q1,size(X,1),1);
     d=abs(X-B)';   % Distance metric
      %scc=repmat(scn,size(Xq,2), 1);
     scc=repmat(sc,size(X,2), 1);
     w=exp(-sqrt(sum((d.^2)./scc.^2)));          %Gaussian Kernal function
     wX = X.*(w'*ones(1,nx));                 % Weighted inputs
     yq(i) = q1*pinv(wX'*X)*wX'*Y;
     unc(i)=w*w'*noisevar;
     %  scc=repmat(sc,size(Xq,1), 1);
%     w=exp(-(d.^2)/sc.^2);          % Gaussian Kernal function
%     wX = X.*(w*ones(1,nx));      % Weighted inputs
 %   waitbar(i/iters,h)
end
%close(h)


load('lungcom');
l=[1, 2, 3, 6, 9, 10, 12, 18];
[row, col]=size(y(l,:));
for i=8:11;
q=isnan(s(:,i));
u=find(q==0);
z=find(u>20);
ytry=u(z);
h_KR=3.5;
a=[1:col];
V=diag(ones(20,1));
for j=1:size(ytry)
    b=setdiff(a,i);
    [ys,ym,yst]=zscore(y');
    XT=ys(b,l);
    XQ=ys(i,l);
    YT=ys(b,ytry(j));
    YQ=ys(i,ytry(j));
    p=1;
    [YHAT1, Var_KR4] = lwr_mm(XT,YT,h_KR,XQ);
    ypred(j,i)=unscore(YHAT1, ym(ytry(j)), yst(ytry(j)));
    Var_kr(j,i)=Var_KR4;
    Var_lwr(j,i)=unscore(Var_KR4, ym(ytry(j)), yst(ytry(j)));
    yact=y(ytry(j),i);
    prederror(j,i)=abs(ypred(j)-yact);
    predper(j,i)=((abs(yact-ypred(j)))./yact).*100;

end
figure
o=union(l,u);
ytrue=y(o,i);
```

```
siztrue=size(ytrue);
g=y(l,i);
nonzp=nonzeros(ypred(:,i));
ytest=[g; nonzp];
tr=union(l, ytry);
nonzv=nonzeros(Var_lwr(:,i));
CI=2*sqrt(nonzv);
hold on
plot(tr-1, ytest, 'k*:')
hold on
plot(o-1, ytrue, 'b.--')
errorbar(ytry-1, nonzp, CI, 'r+');
legend('Predicted', 'Actual', '95% Confidence Interval');
xlabel('Elapsed Days');
ylabel('Volume (cc)');
title(['Patient ' num2str(i) ])
end

function [mse ypred yact unc traindat realactpreddat realpreddat YQ] =
lunggetgraph(org, x, sc, gax, nodays, novar)
[row col]=size(x);
[row col]=size(x);
days=sort(unique(round(gax(1:nodays)))); %days used for observation (remember
that l=1 goes w/ day=0)
ypred=zeros(row,1);
h_KR=gax(nodays+1:col+nodays);
traindat=zeros(row,1);
unc=zeros(row,1);
yact=zeros(row,1);
i=17; %iterate through patients using other patients to predict
a=[1:col]; %used to set difference
b=setdiff(a,i); %set difference to get vector of all patients except the one
that is being predicted
k=find(isnan(org(:,i))==0);
[trainday]=find(k<20);
[testday]=find(k>20);
testdat=k(testday);
    for m=1:(size(testdat)-1)
    b=setdiff(a,i); %set difference to get vector of all patients except the
one that is being predicted
    XT=x([days],b); %set x-training data
    XQ=x([days],i); %set x-query data
    YT=x(testdat(m),b); %set y-training data
    YQ(m)=x(testdat(m),i); %set y-query data
    yact(testdat(m))=x(testdat(m),i); %row 65 has the final volume for each
patient at the end of their treatment
    localsc=h_KR(b);
    [ypred(testdat(m)) unc(testdat(m))] = lwpr1_mm_local(XT',YT',localsc,
XQ', novar); %predict final volume
    end


    realactpreddat=yact.*sc(1,i);
    realpreddat=ypred.*sc(1,i);
    realunc=sqrt(unc.*sc(1,i));
[I J]=find(realpreddat~=0);
```

```
for t=1:size(I)
    error(t)=abs(realpreddat(I(t), J(t))-realactpreddat(I(t), J(t)));
end
figure
plot(k(1:size(k,1)-1),org(k(1:size(k,1)-1),i),'b.')
hold on
plot(x(1:testdat(m),i).*sc(1,i), 'b')
hold on
plot(days, x(days,i).*sc(1,i), 'mo')
errorbar(testdat(1:m), realpreddat(testdat(1:m)), realunc(testdat(1:m)), 'r')
for t=1:size(I)
    error(t)=abs(realpreddat(I(t), J(t))-realactpreddat(I(t), J(t)))
end
testdat
mse=sum(error);

end


function [mse ypred yact unc] = lungscVollocalnopath(x, sc, gax, nodays,
novar)
[row col]=size(x);
days=sort(round(gax(1:nodays))); %days used for observation (remember that
l=1 goes w/ day=0)
h_KR=gax(nodays+1:col+nodays);
for i=1:size(x,2) %iterate through patients using other patients to predict
yact(i)=x(65,i); %row 65 has the final volume for each patient at the end of
their treatment
a=[1:col]; %used to set difference
b=setdiff(a,i); %set difference to get vector of all patients except the one
that is being predicted
XT=x([days],b); %set x-training data
XQ=x([days],i); %set x-query data
YT=x(65,b); %set y-testing data
YQ=x(65,i); %set y-query data
localsc=h_KR(b);
p=1;
[ypred(i) unc(i)] = lwpr1_mm_local(XT',YT',localsc, XQ', novar); %predict
final volume
end
mse=100*mean(abs(yact(1:col)-ypred(1:col)))+mean(abs(yact(1:col).*sc(1:col)-
ypred(1:col).*sc(1:col)));

% prederror(B,:)=abs(ypred(B,:)-yact);
% predper(B,:)=((abs(yact-ypred(B,:))./yact).*100;

end


function [mse ypred yact unc] = lungscmasslocalnopath(x, sc, gax, novar)
[row col]=size(x);
days=[1 2 3 4 7 8 9 10 11 12 13 14 15 16 17 19 20]; %days used for
observation (remember that l=1 goes w/ day=0)
h_KR=gax(1:col);
for i=1:size(x,2) %iterate through patients using other patients to predict
```

```
yact(i)=x(65,i); %row 65 has the final volume for each patient at the end of
their treatment
a=[1:col]; %used to set difference
b=setdiff(a,i); %set difference to get vector of all patients except the one
that is being predicted
XT=x([days],b); %set x-training data
XQ=x([days],i); %set x-query data
YT=x(65,b); %set y-testing data
YQ=x(65,i); %set y-query data
localsc=h_KR(b);
p=1;
[ypred(i) unc(i)] = lwpr1_mm_local(XT',YT',localsc, XQ', novar); %predict
final volume
end
mse=80*mean(abs(yact(1:col)-ypred(1:col)))+(mean(abs(yact(1:col).*sc(1:col)-
ypred(1:col).*sc(1:col))));

% prederror(B,:)=abs(ypred(B,:)-yact);
% predper(B,:)=((abs(yact-ypred(B,:)))./yact).*100;

end


function [yq unc] = lwpr1_mm_local(X,Y,sc,q, noisevar)
%
%   Locally Weighted Polynomial Regression with Local Gaussian kernal and
Euclidean distance.
%   Multivariate input matrix.
%
% X     - input data matrix 19X10
% Y     - output data matrix
% sc    - bandwidth vector
% q     - matrix of query vectors: row = observation, column = inputs
%
% yq     - prediction for q

% Hines 2002
% University of Tennessee
X=[X X.^2]; % compute squares of inputs
q=[q q.^2];
%[X xm xs]=zscore1(X);
%q=zscore1(q,xm,xs);

[n,nx] = size(X);           % n is number of patterns
[n,ny] = size(Y);         % nx and ny are number of inputs
                            % and outputs

iters=size(q,1);
yq=zeros(1,iters);
%h = waitbar(0,'Please wait...');

Xq=X;                       % Original Inputs
X = [X ones(n,1)];          % Inputs appended with ones for intercept
nx = nx+1;                  % New number of inputs
scn=[sc sc.^2];
```

```
scn=[scn 1];
for i=1:iters;
     q1 = [q(i,:) 1]; % Query appended with ones
     B=repmat(q1,size(X,1),1);
     d=abs(X-B)';    % Distance metric
      %scc=repmat(scn,size(Xq,2), 1);
     scc=repmat(sc,size(X,2), 1);
     w=exp(-sqrt(sum((d.^2)./scc.^2)));         %Gaussian Kernal function
     wX = X.*(w'*ones(1,nx));                % Weighted inputs
     yq(i) = q1*pinv(wX'*X)*wX'*Y;
     unc(i)=w*w'*noisevar;
     %  scc=repmat(sc,size(Xq,1), 1);
%     w=exp(-(d.^2)/sc.^2);          % Gaussian Kernal function
%     wX = X.*(w*ones(1,nx));      % Weighted inputs
 %   waitbar(i/iters,h)
end
%close(h)


function [mse ypred yact unc] = lungscVolnoday(x, sc, gax, nodays, novar)
[row col]=size(x);
days=sort(unique(round(gax(1:nodays)))) %days used for observation (remember
that l=1 goes w/ day=0)
h_KR=gax(nodays+1:col+nodays);
for i=1:size(x,2) %iterate through patients using other patients to predict
yact(i)=x(65,i); %row 65 has the final volume for each patient at the end of
their treatment
a=[1:col]; %used to set difference
b=setdiff(a,i); %set difference to get vector of all patients except the one
that is being predicted
XT=x([days],b); %set x-training data
XQ=x([days],i); %set x-query data
YT=x(65,b); %set y-testing data
YQ=x(65,i); %set y-query data
localsc=h_KR(b);
p=1;
[ypred(i) unc(i)] = lwpr1_mm_local(XT',YT',localsc, XQ', novar); %predict
final volume
end
mse=100*mean(abs(yact(1:col)-ypred(1:col)))+mean(abs(yact(1:col).*sc(1:col)-
ypred(1:col).*sc(1:col)));

% prederror(B,:)=abs(ypred(B,:)-yact);
% predper(B,:)=((abs(yact-ypred(B,:))./yact).*100;

End


function plotaakr(p, act, numst, numend, myavi)
x=[1:1:size(p,1)];
count=0;
close
for i=numst:1:numend
count=count+1
plot(x,act(:,i),'.',x,p(:,i),'k','LineWidth',1.5);
ylabel('Detector Count');
```

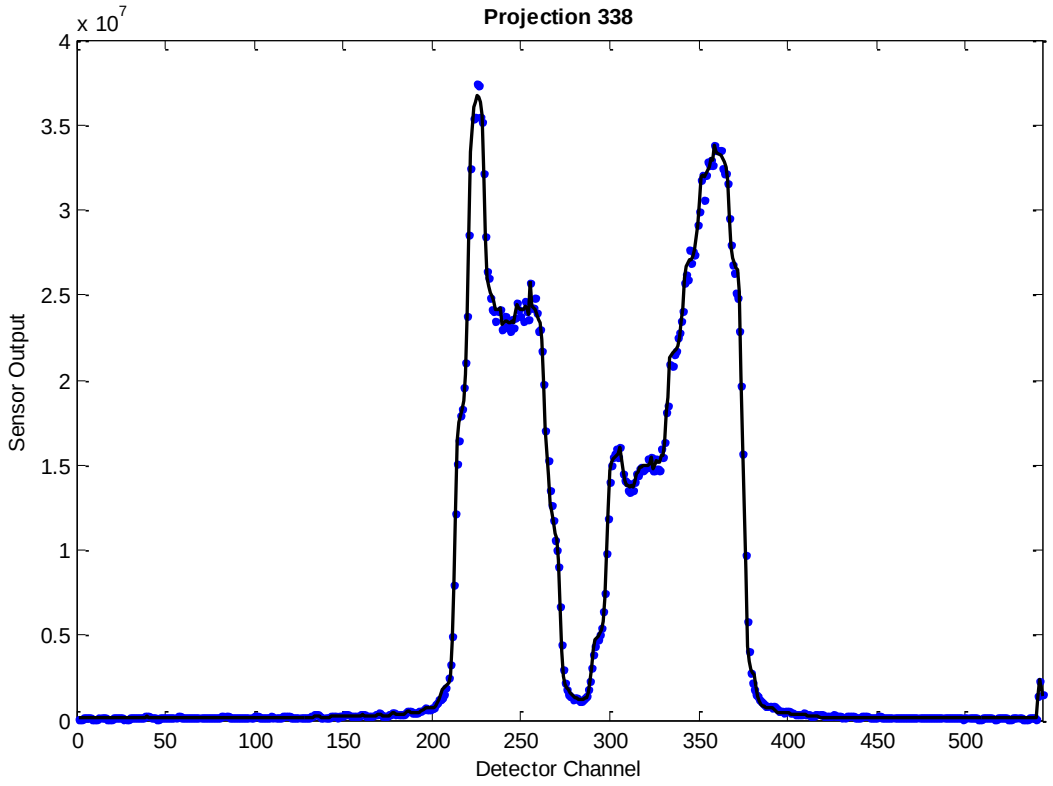
```
xlabel('Detector Channel');
legend('Actual','Predicted','Location','Best');
frame = getframe(gca, [0 0 435 343]);
myavi = addframe(myavi,frame);
end
end

function [errormap, detecteders]=mlcerrordet(fwhmtable1, fwhmtable2, frames,
threshold);
for i=1:size(fwhmtable1,1)
for j=1:size(fwhmtable1,2)
if ((fwhmtable1(i,j)==0) | (fwhmtable2(i,j)==0))
errormap(i,j)=0;
else
errormap(i,j)=((fwhmtable1(i,j)/frames)-(fwhmtable2(i,j)/frames))*100;
end
end
end

[r c]=find(abs(errormap)>threshold);
k=unique(r);
count=1;
for i=1:size(k)
t=find(r==k(i));
if size(t)==1;
newerror(count)=k(i);
count=count+1;
end
end

j=[r c];
for k=1:size(newerror,2)
l=find(r==newerror(k));
newj(k,:)=j(l,:);
end

detecteders=unique(newj, 'rows');

end


function [meanmse yq ya yact yt] = lunglpr(x,y,nodays)
days=sort(round(x(1:nodays)))
[nox1, nox2]=size(x);
k=x(nodays+1:nox2); %h_kr
clear ypred
clear sc
gd=[days, 59];
[ys,ym,yst]=zscore(y(gd,:)');
xdat=ys(:,1:nodays);
ydat=ys(:,nodays+1);
for i=1:20
     %row 59 has the final volume for each patient at the end of their
treatment
    a=[1:size(y,2)]; %used to set difference
```

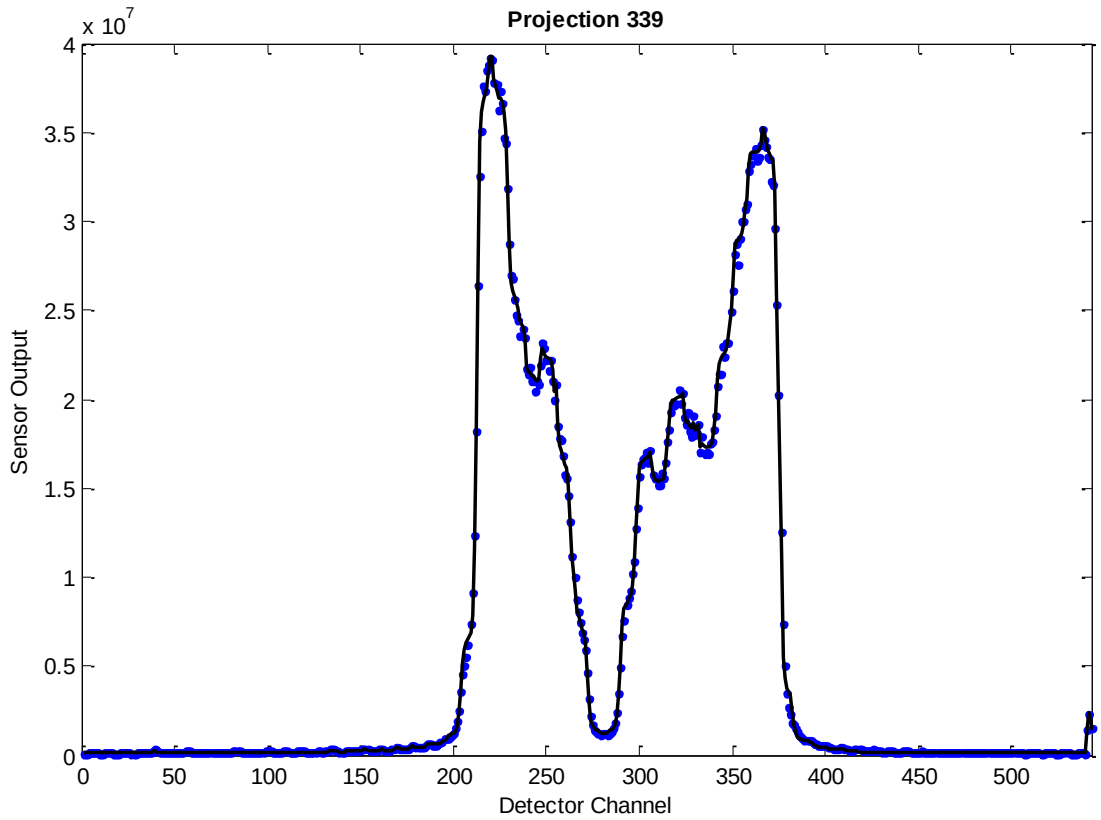
```
    b=setdiff(a,i); %set difference to get vector of all patients except the
one that is being predicted
    sc=k(b);
    XT=xdat(b,:);
    YT=ydat(b,:);
    XQ=xdat(i,:);
    yq(i) = lwpr1_mm_local(XT,YT,sc,XQ);
%     yp(i)=lwpr_mm(XT,YT,3,XQ);
%     yl(i)=lwr_mm(XT,YT,3,XQ);
   yact(i)=ys(i, nodays+1);
   yt(i)=unscore(yq(i), ym(nodays+1), yst(nodays+1));
%     yc(i)=unscore(yp(i), ym, ystd);
%     yv(i)=unscore(yl(i), ym, ystd);
     ya(i)=unscore(yact(i), ym(nodays+1), yst(nodays+1));
end
meanmse=(sumsqr(yt-ya))/length(ya)


i=setdiff(1:20, 3);
for j=1:size(i,2);
fid = fopen(['F:\MatLab Sinogram Data\Prostate Test\       ', num2str(i(j)),
'.sin']);
if fid==-1
    fid = fopen(['F:\MatLab Sinogram Data\Prostate Test\      ',
num2str(i(j)), '.sin']);
end
if fid~=-1
sinogram(j).data=fread(fid, [643 757], 'float');
sinogram(j).trial=i(j);
keeptrack(j)=i(j);
end
end
count=0
for b=1:size(i,2)-1
    for k=b+1:size(i,2)
        count=count+1
        sinogram(count).subtraction=sinogram(i).data(38:597,:)-
sinogram(j).data(38:597,:);
        sinogram(count).subtractkeeptrack=[i,j];
    end
end
```

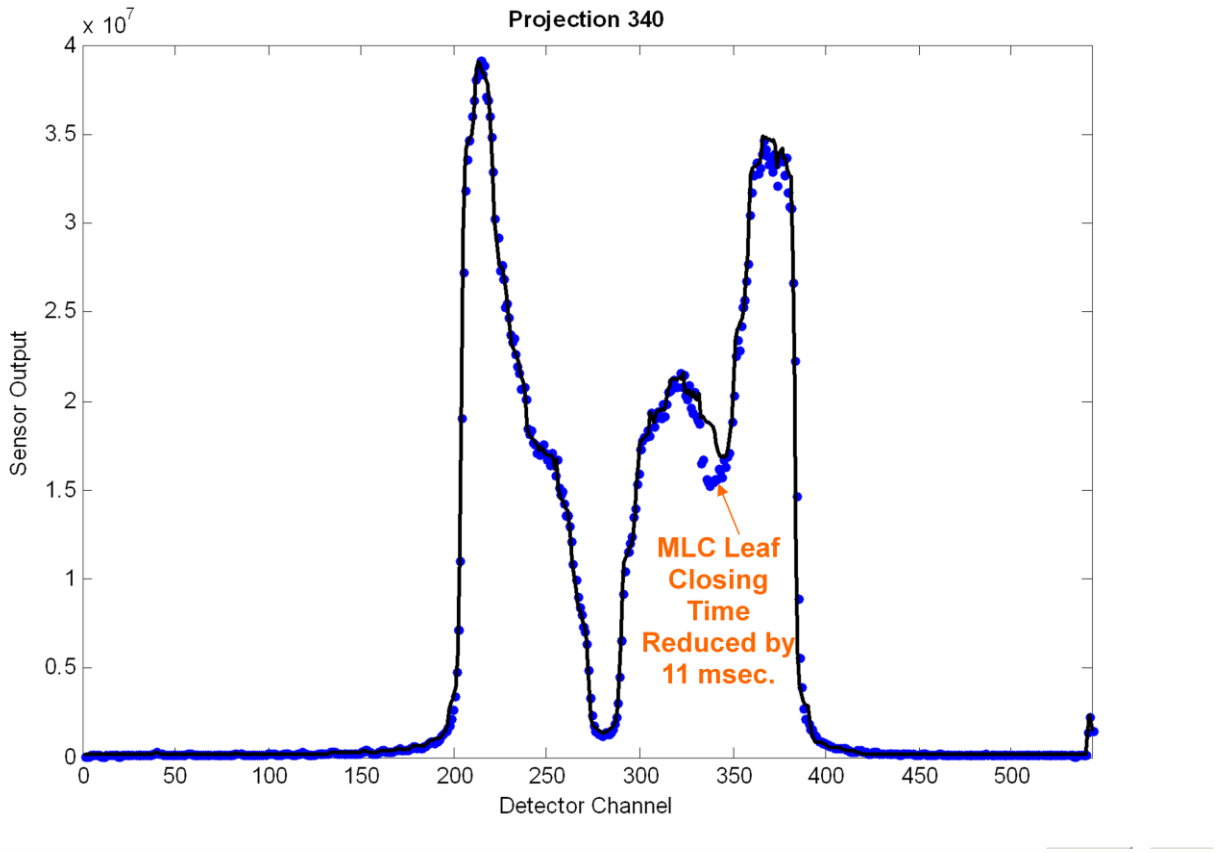
APPENDIX C. AAKR DETECTION OF SINOGRAM ERRORS

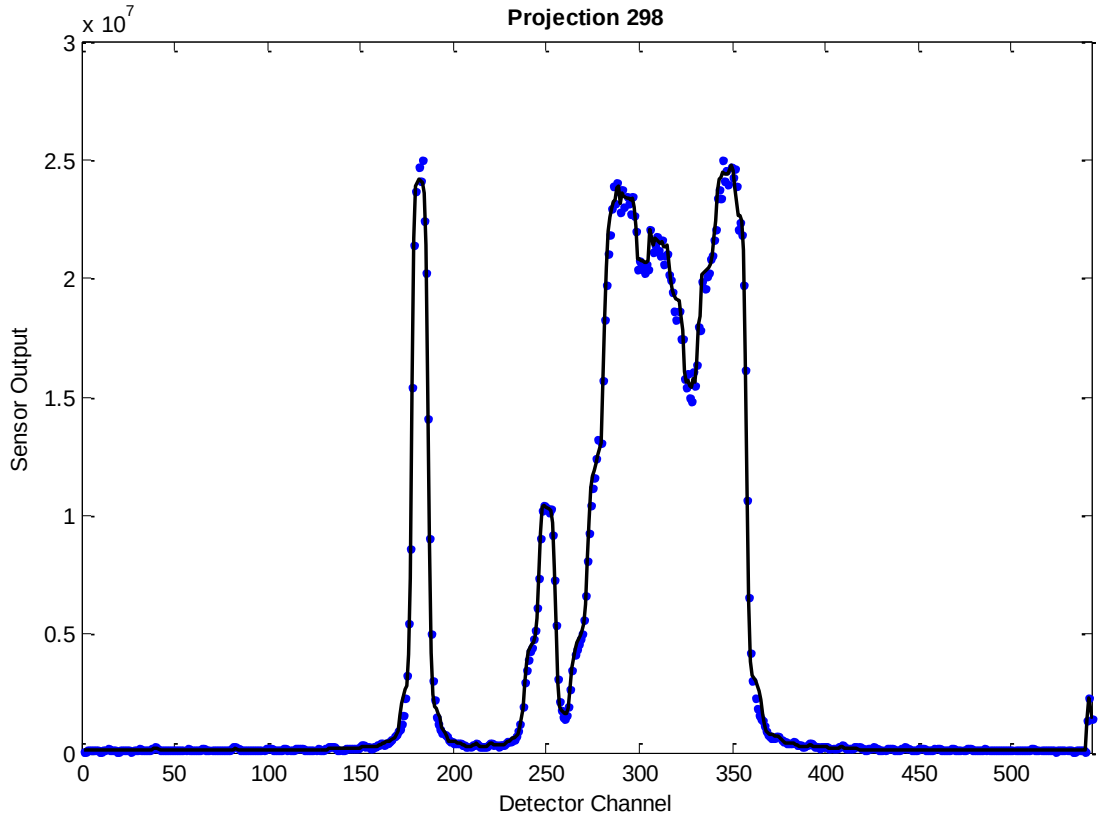
This appendix presents more figures of the AAKR model detecting MLC positional errors. The model predictions for the unmodified projections which are prequel to the faulted projection are also shown to help illustrate how the model responds under normal operating conditions. In all of the figures the blue dots are the measured detector counts and the black lines are the AAKR predictions.

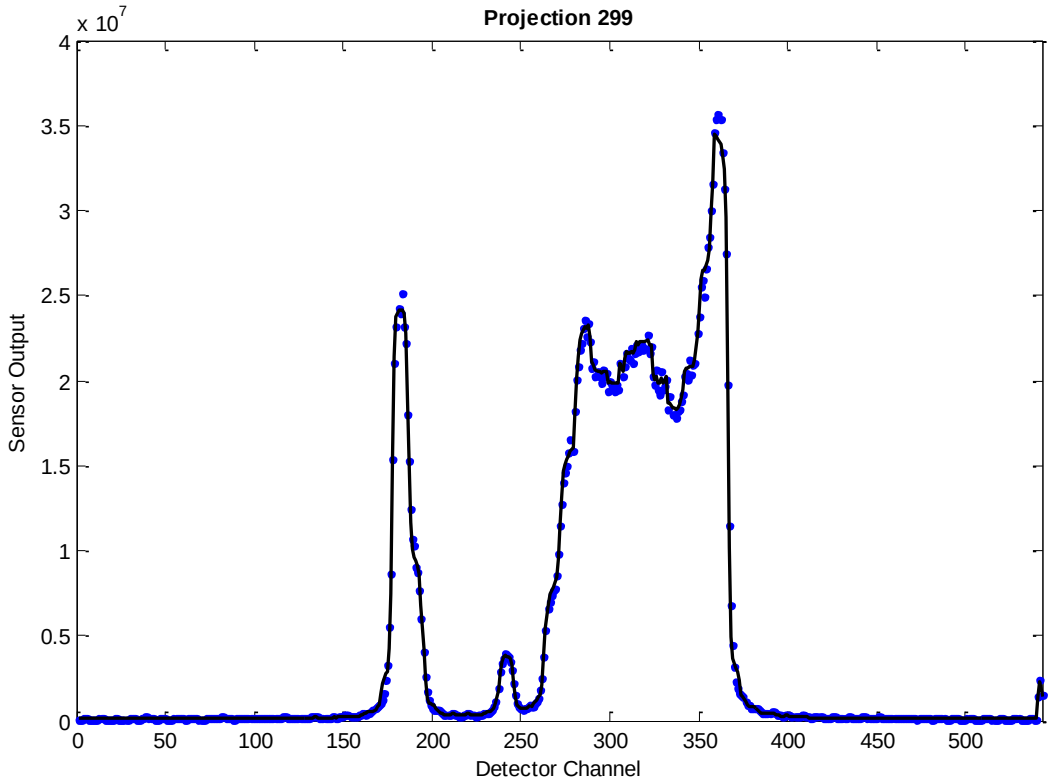
HEAD & NECK DELIVERY AAKR PREDICTIONS

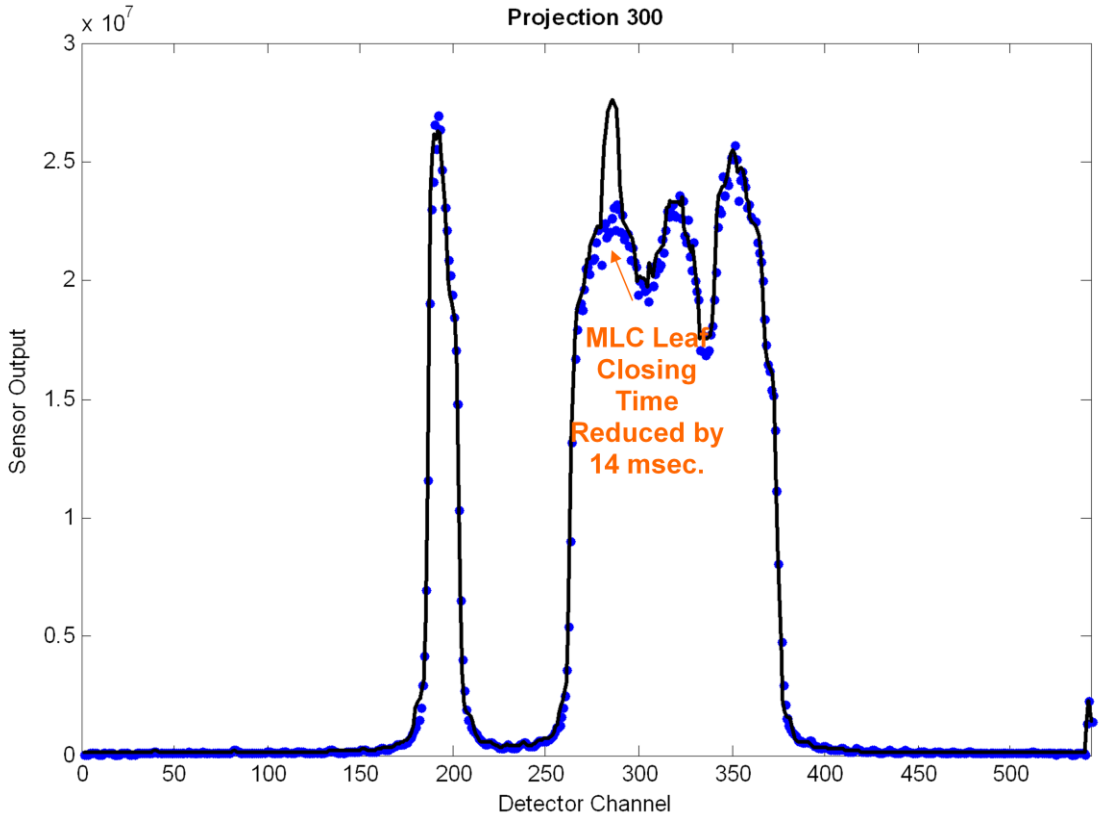


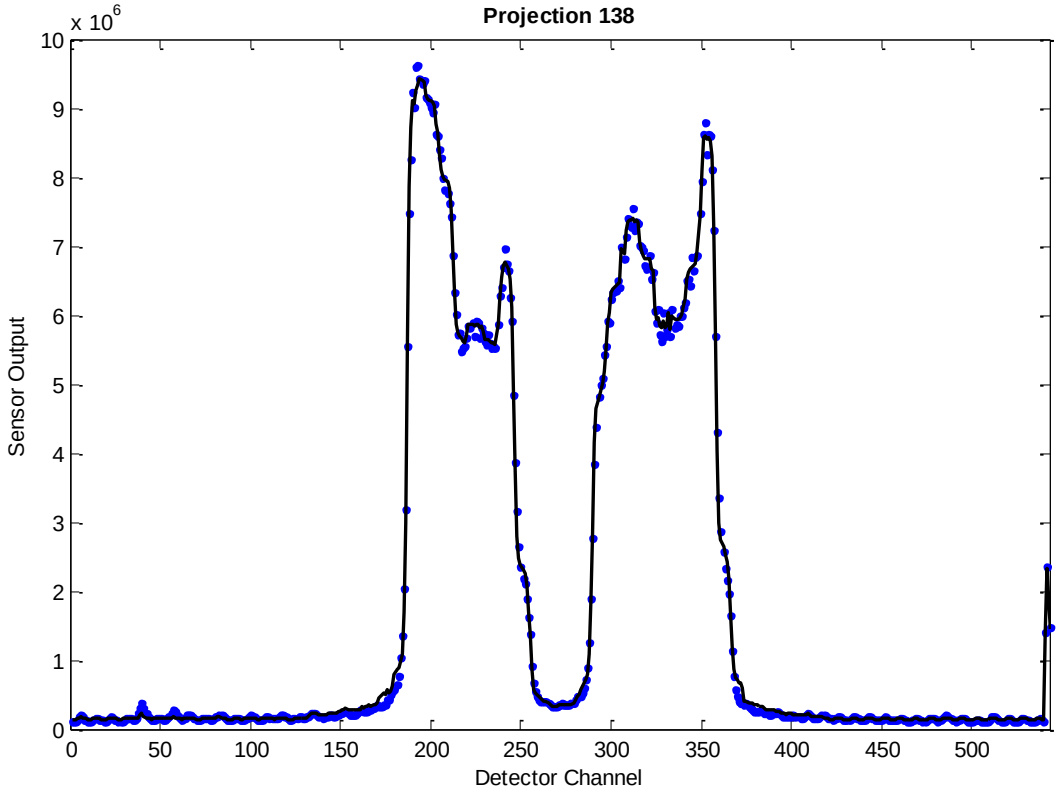


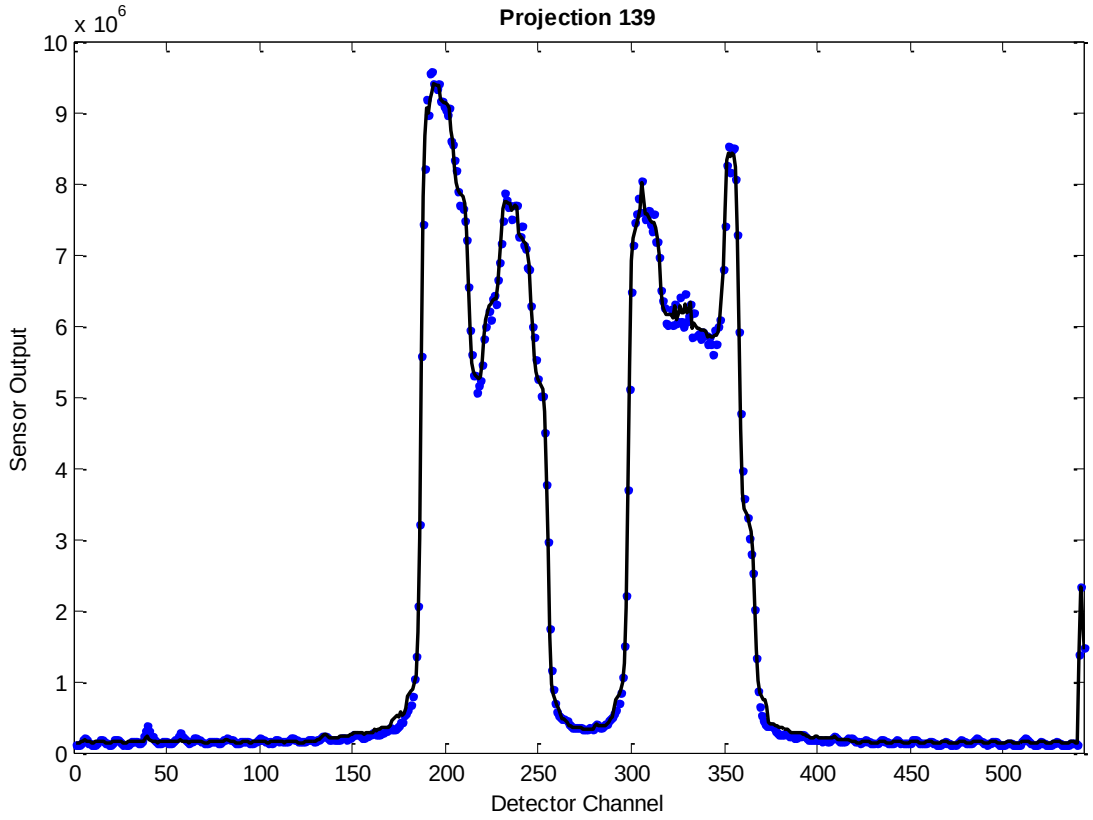

Projection 340
x 10^7
Sensor Output
Detector Channel
MLC Leaf Closing Time Reduced by 11 msec.
0
0.5
1
1.5
2
2.5
3
3.5
4
0
50
100
150
200
250
300
350
400
450
500

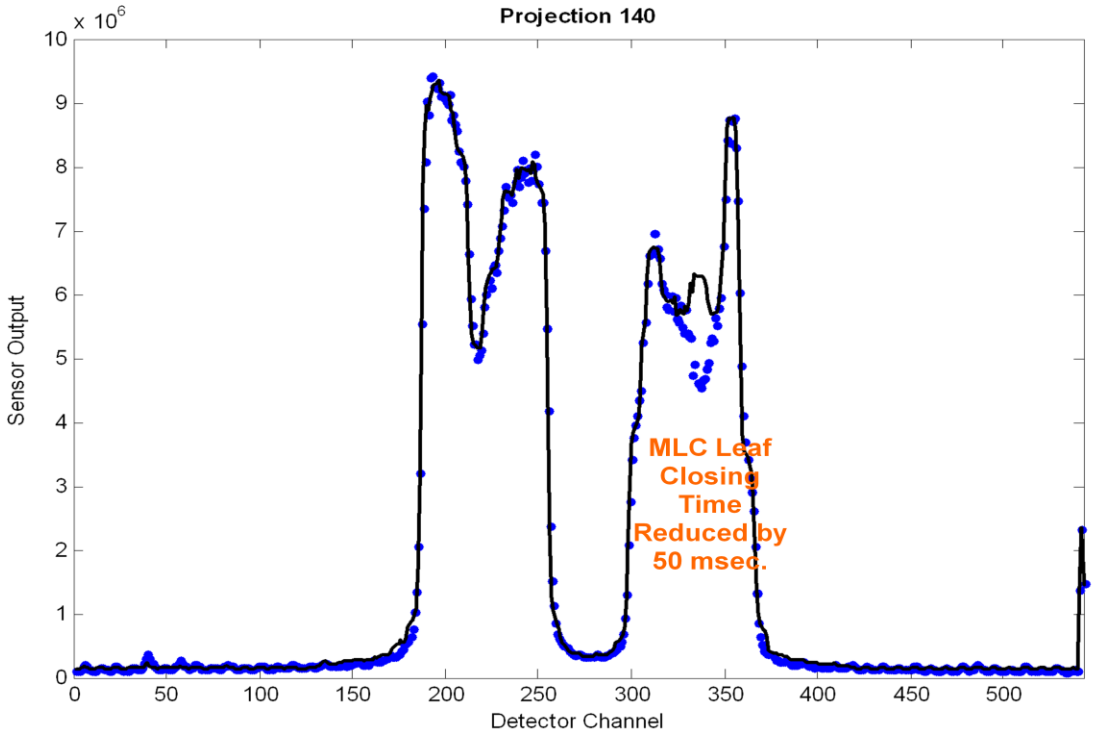

Projection 298
x 10^7
Sensor Output
Detector Channel
0
0.5
1
1.5
2
2.5
3
0
50
100
150
200
250
300
350
400
450
500


Projection 299
x 10^7
Sensor Output
Detector Channel
0
0.5
1
1.5
2
2.5
3
3.5
4
0
50
100
150
200
250
300
350
400
450
500

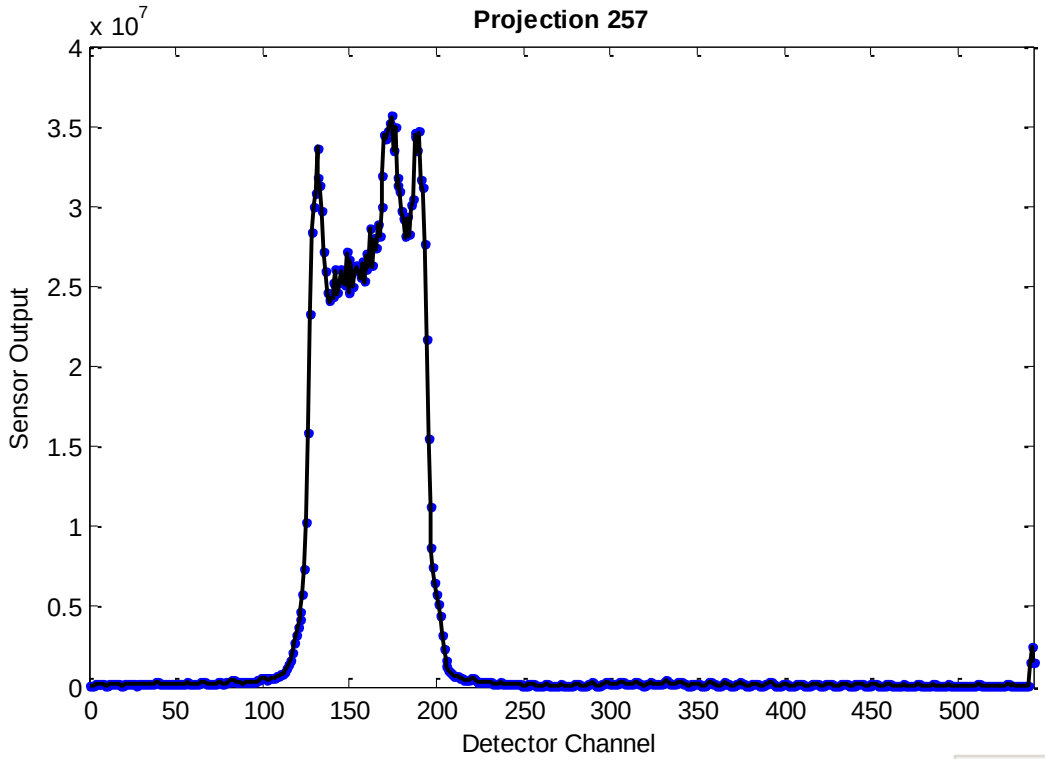

Projection 300
x 10^7
Sensor Output
Detector Channel
MLC Leaf Closing Time Reduced by 14 msec.

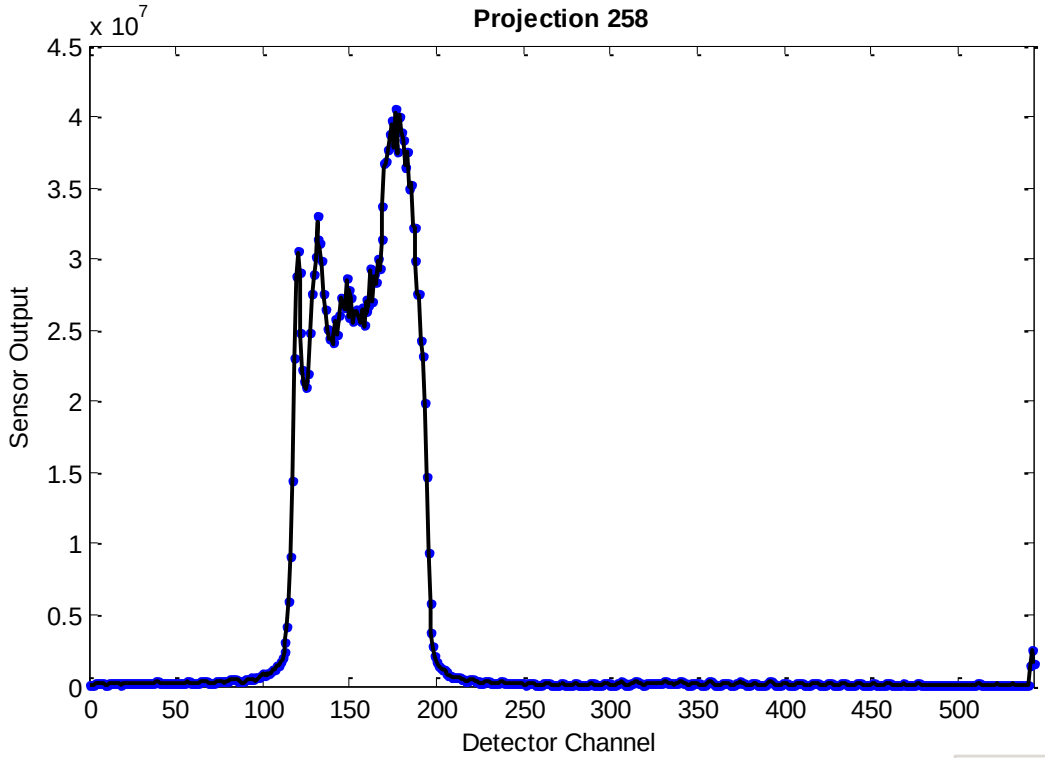

Projection 138
x 10^6
Sensor Output
Detector Channel

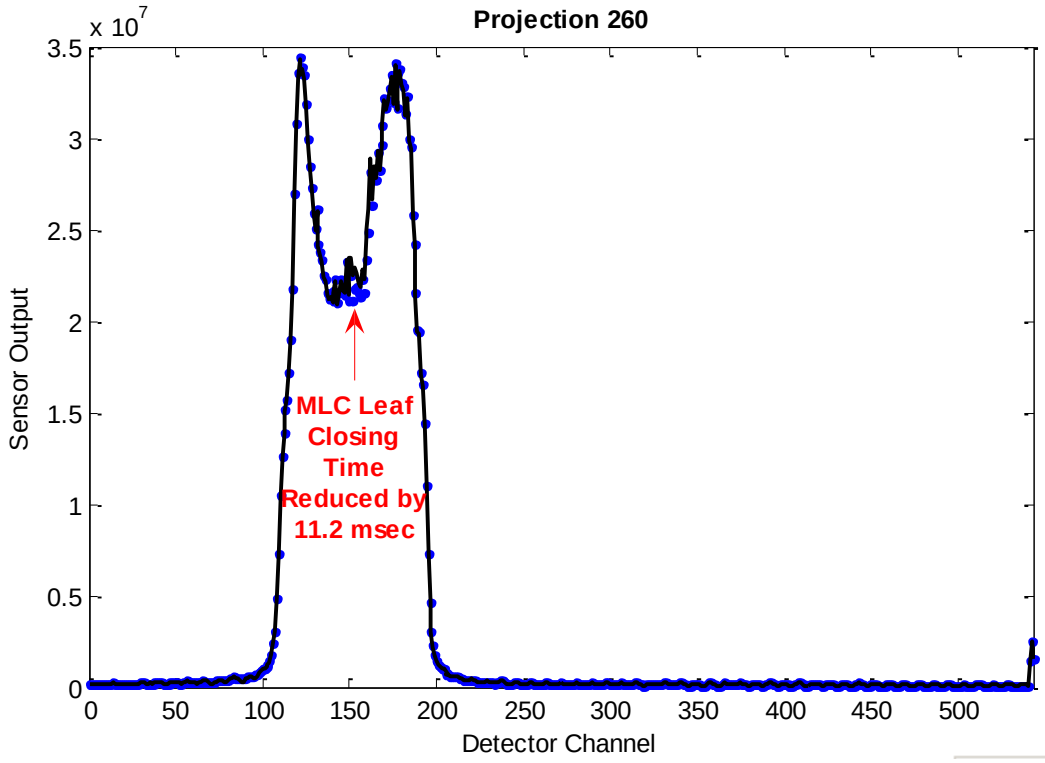

Projection 139
x 10^6
Sensor Output
Detector Channel
0
1
2
3
4
5
6
7
8
9
10
0
50
100
150
200
250
300
350
400
450
500

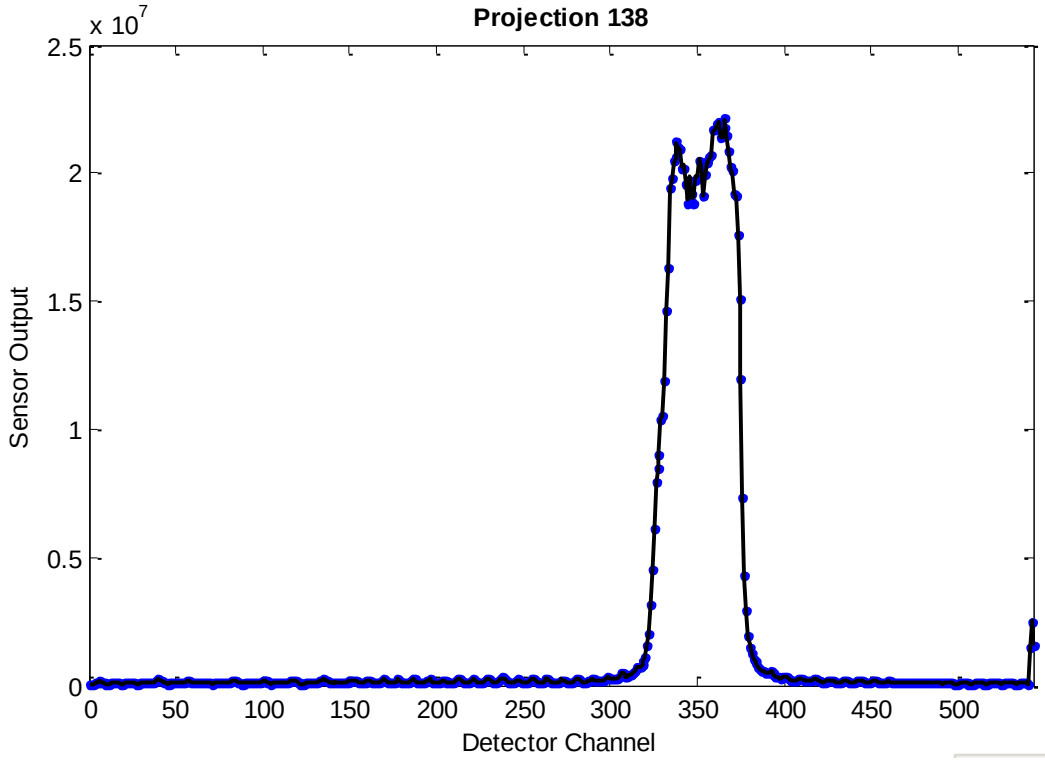

Projection 140
x 10^6
Sensor Output
Detector Channel
MLC Leaf Closing Time Reduced by 50 msec.
0
1
2
3
4
5
6
7
8
9
10
0
50
100
150
200
250
300
350
400
450
500

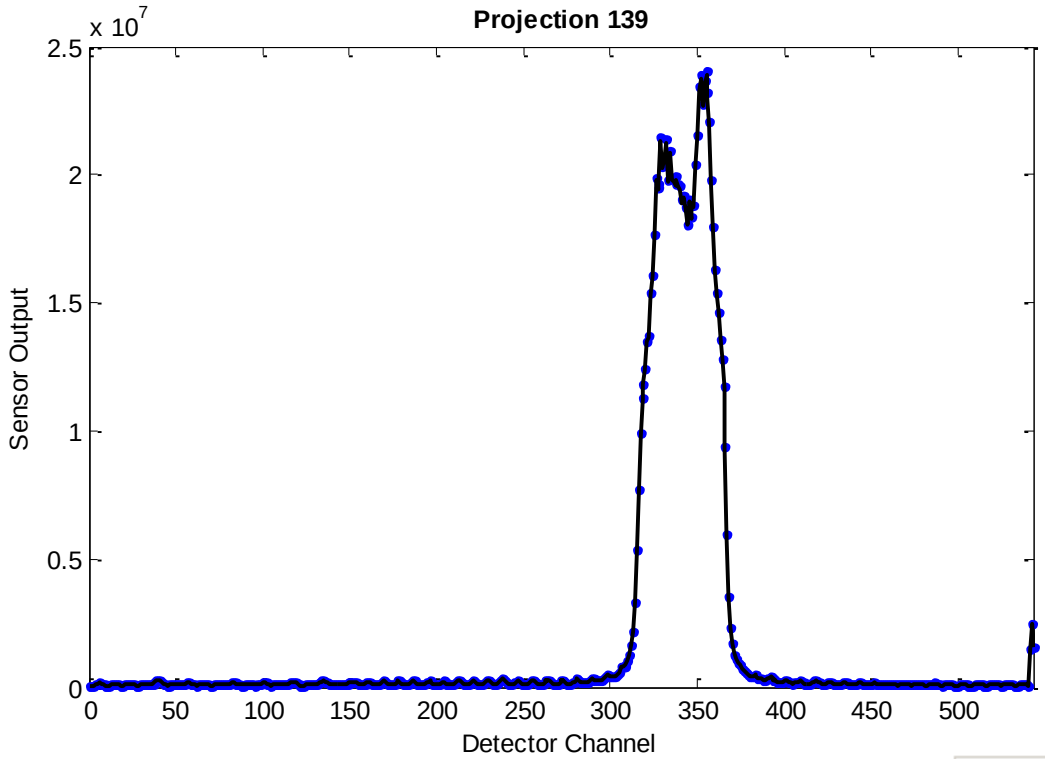
LUNG DELIVERY AAKR PREDICTIONS

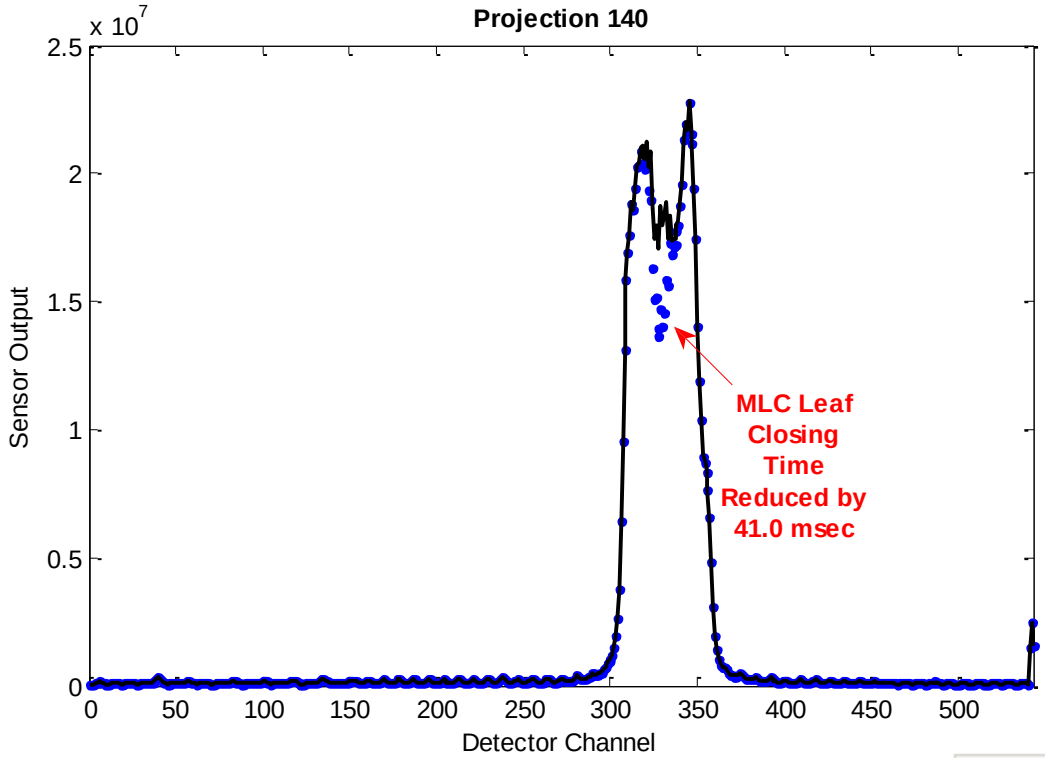


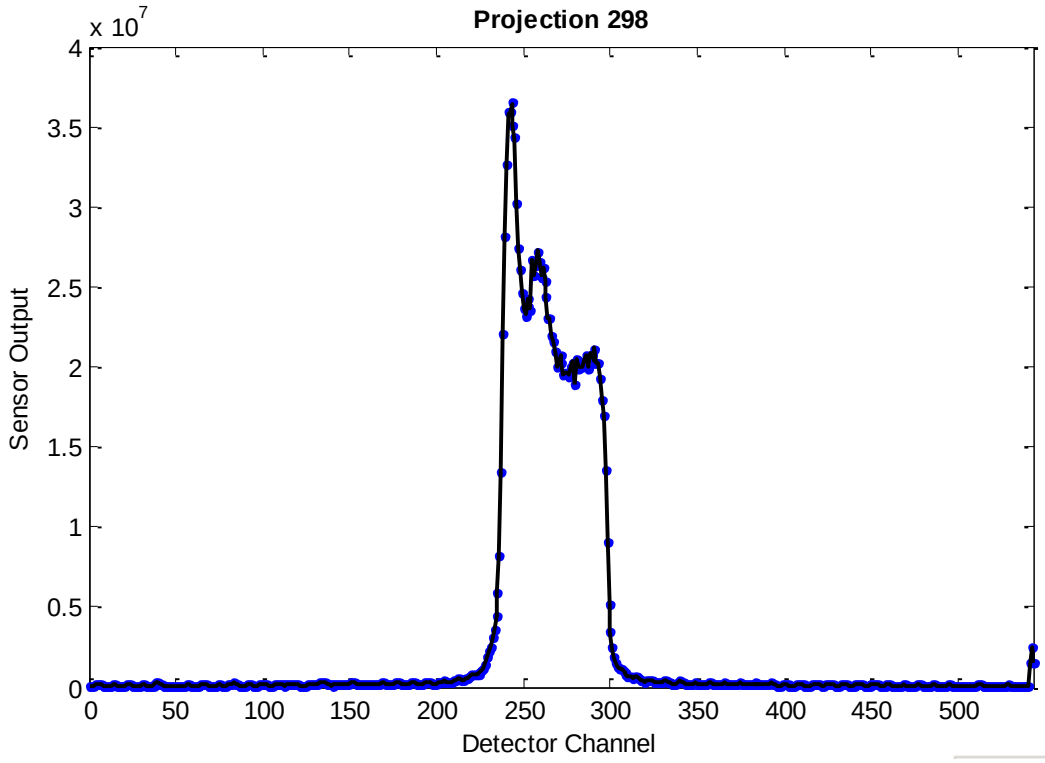


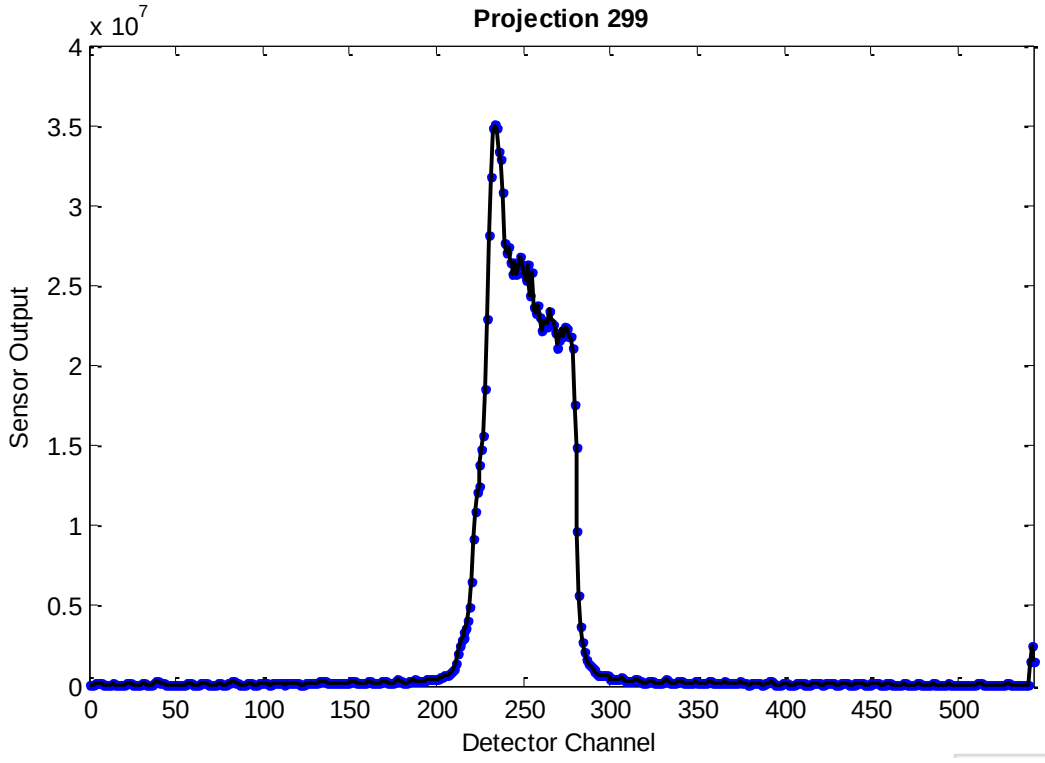

Projection 260
x 10^7
Sensor Output
Detector Channel
MLC Leaf Closing Time Reduced by 11.2 msec

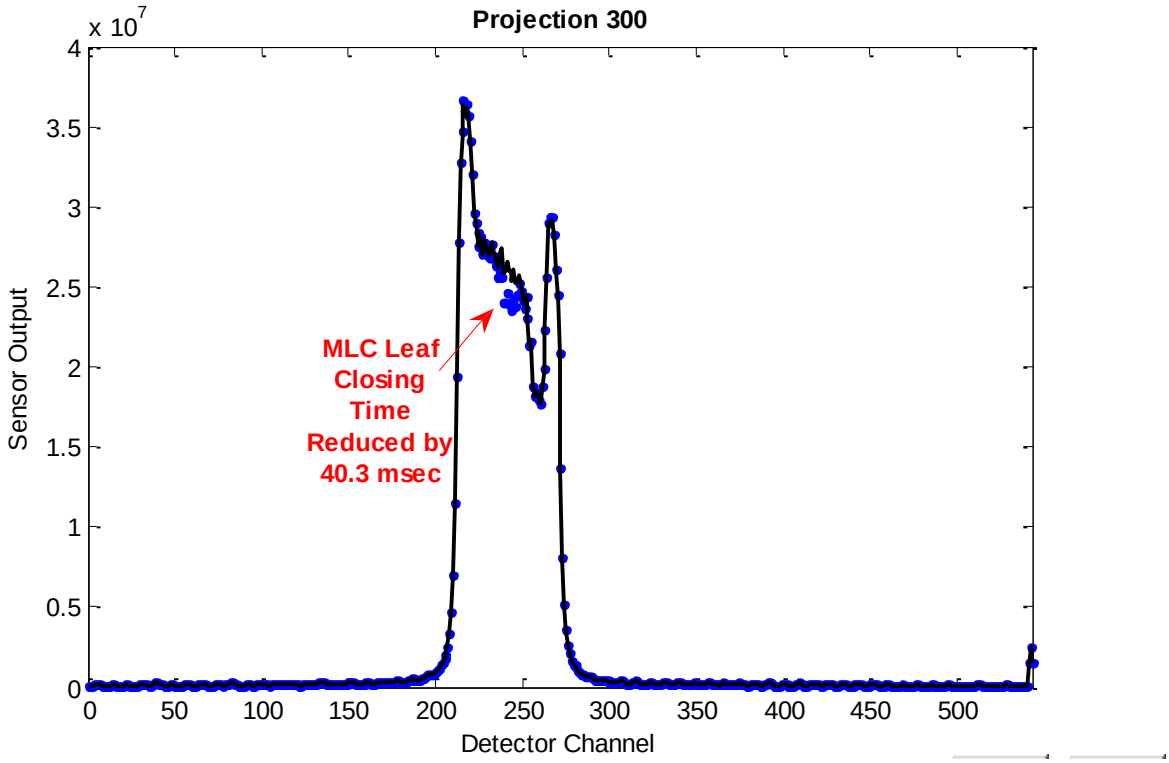

Projection 138
x 10^7
Sensor Output
Detector Channel


Projection 139
x 10^7
2.5
2
1.5
1
0.5
0
Sensor Output
0
50
100
150
200
250
300
350
400
450
500
Detector Channel

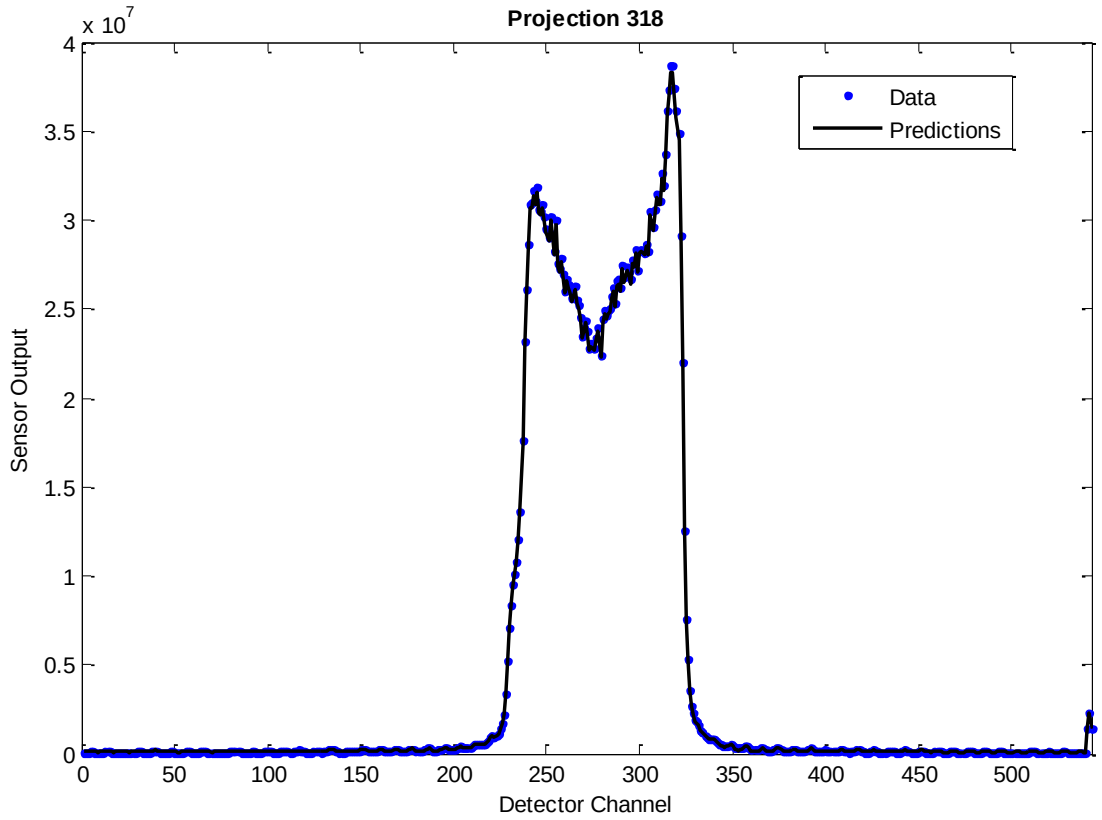

Projection 140
x 10^7
2.5
2
1.5
1
0.5
0
Sensor Output
MLC Leaf Closing Time Reduced by 41.0 msec
0
50
100
150
200
250
300
350
400
450
500
Detector Channel

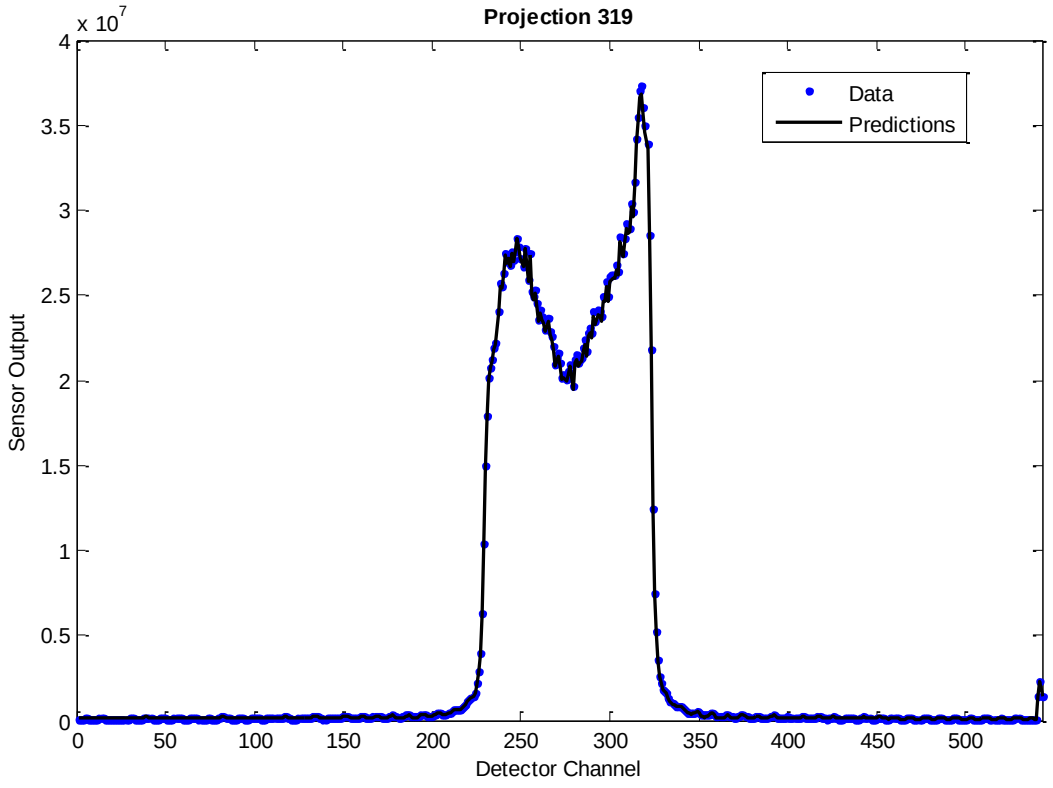

Projection 298
x 10^7
Sensor Output
Detector Channel
0
0.5
1
1.5
2
2.5
3
3.5
4
0
50
100
150
200
250
300
350
400
450
500

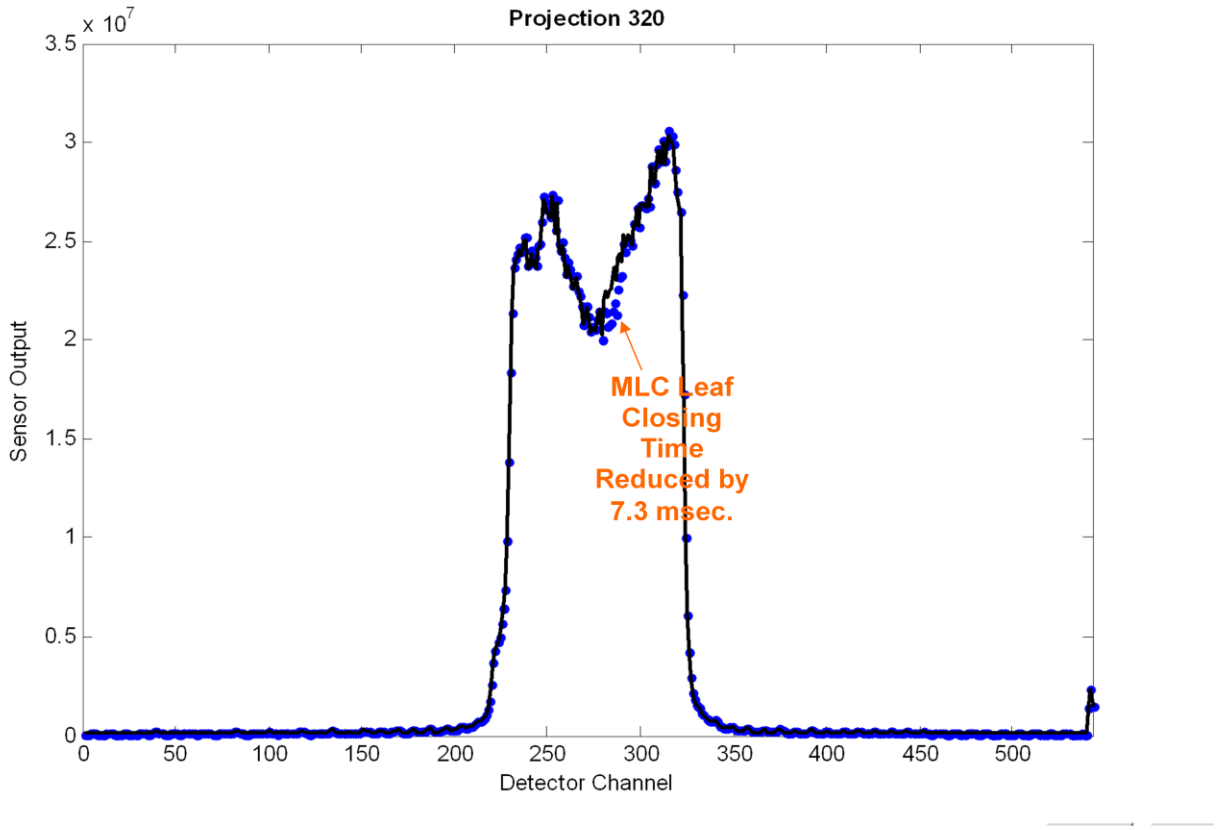

Projection 299
x 10^7
Sensor Output
Detector Channel
0
0.5
1
1.5
2
2.5
3
3.5
4
0
50
100
150
200
250
300
350
400
450
500

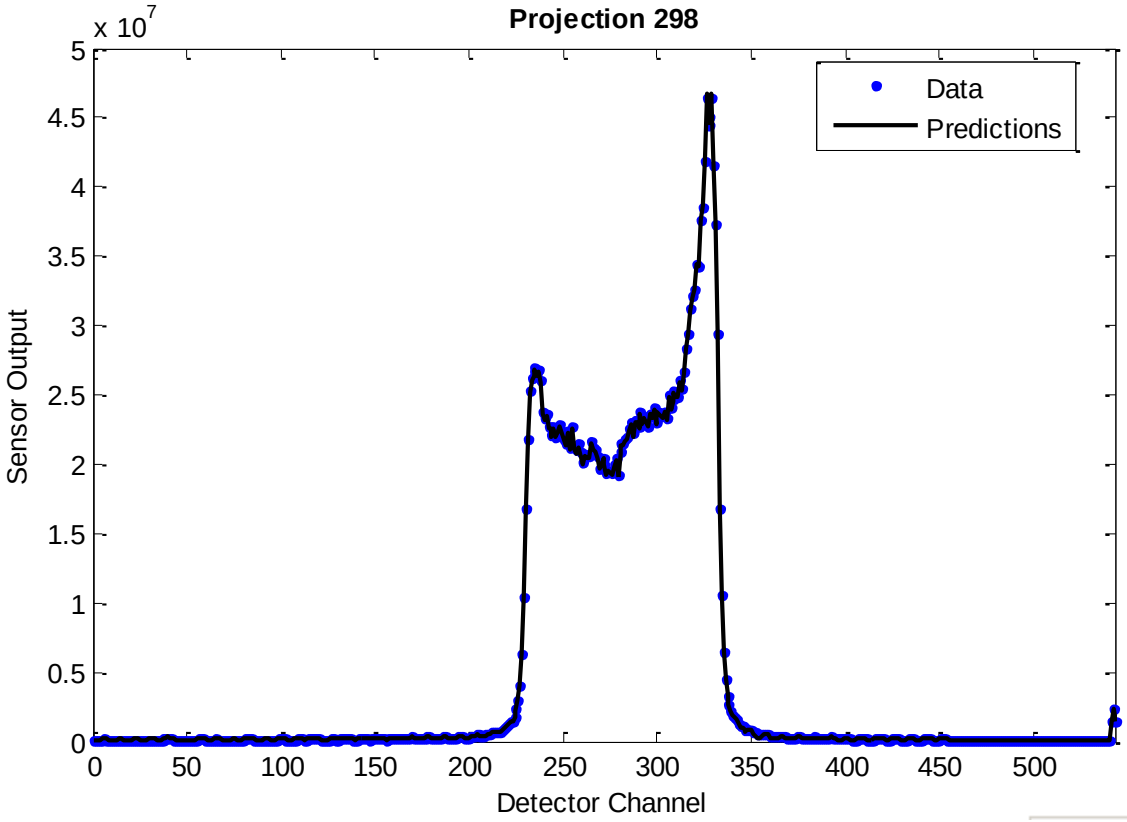

Projection 300
x 10^7
Sensor Output
Detector Channel
MLC Leaf Closing Time Reduced by 40.3 msec
0
0.5
1
1.5
2
2.5
3
3.5
4
0
50
100
150
200
250
300
350
400
450
500

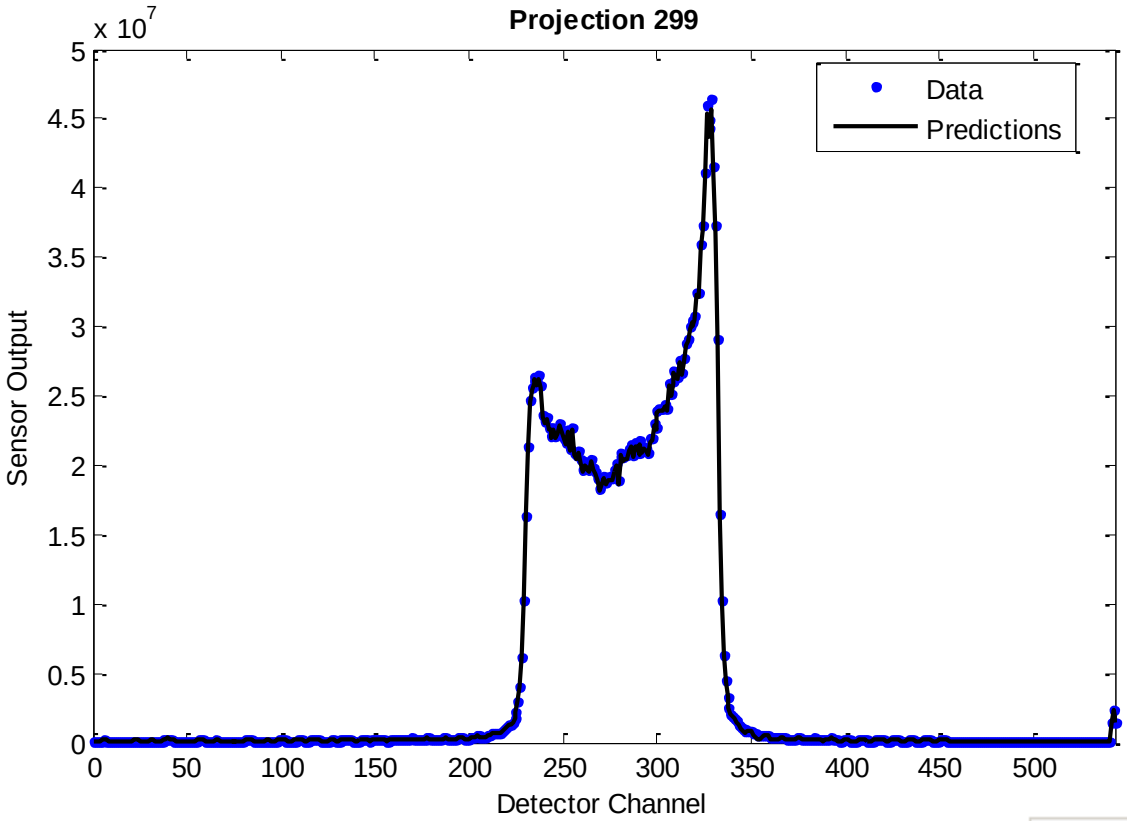
PROSTATE DELIVERY AAKR PREDICTIONS

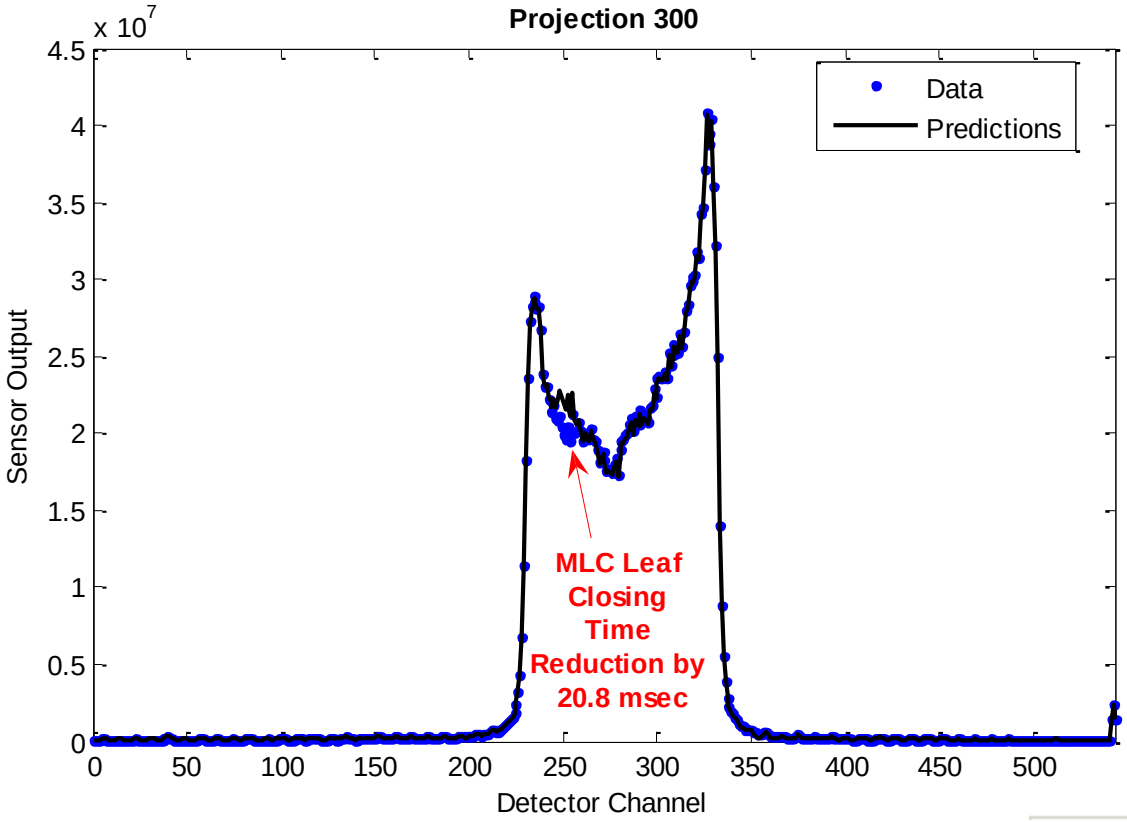


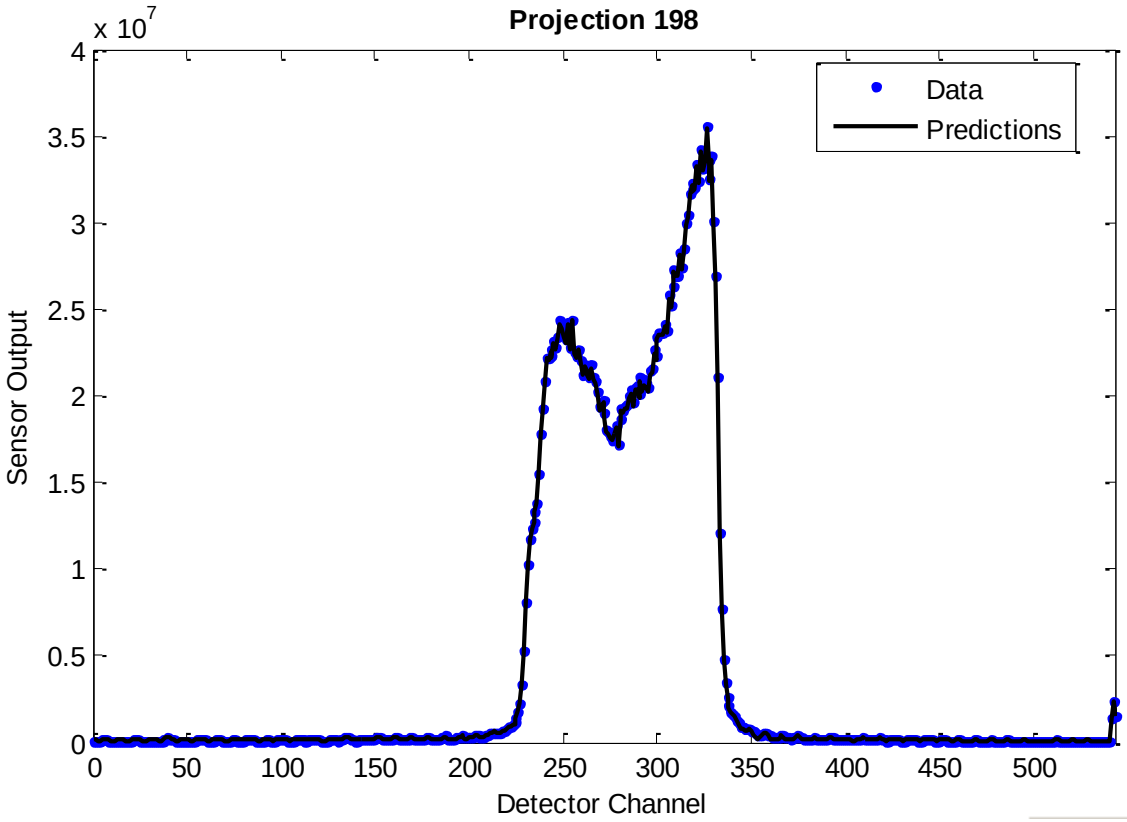


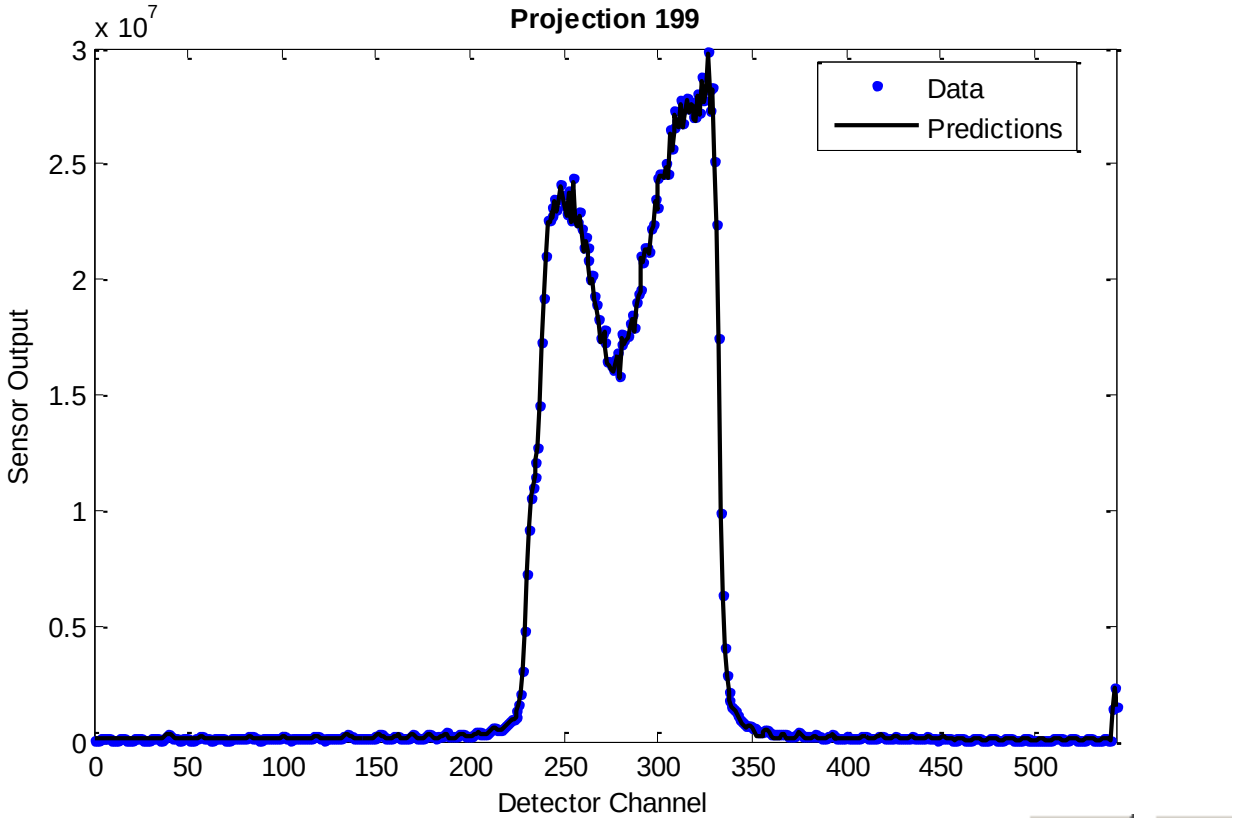

Projection 320
x 10^7
3.5
3
2.5
2
1.5
1
0.5
0
Sensor Output
0
50
100
150
200
250
300
350
400
450
500
Detector Channel
MLC Leaf Closing Time Reduced by 7.3 msec.

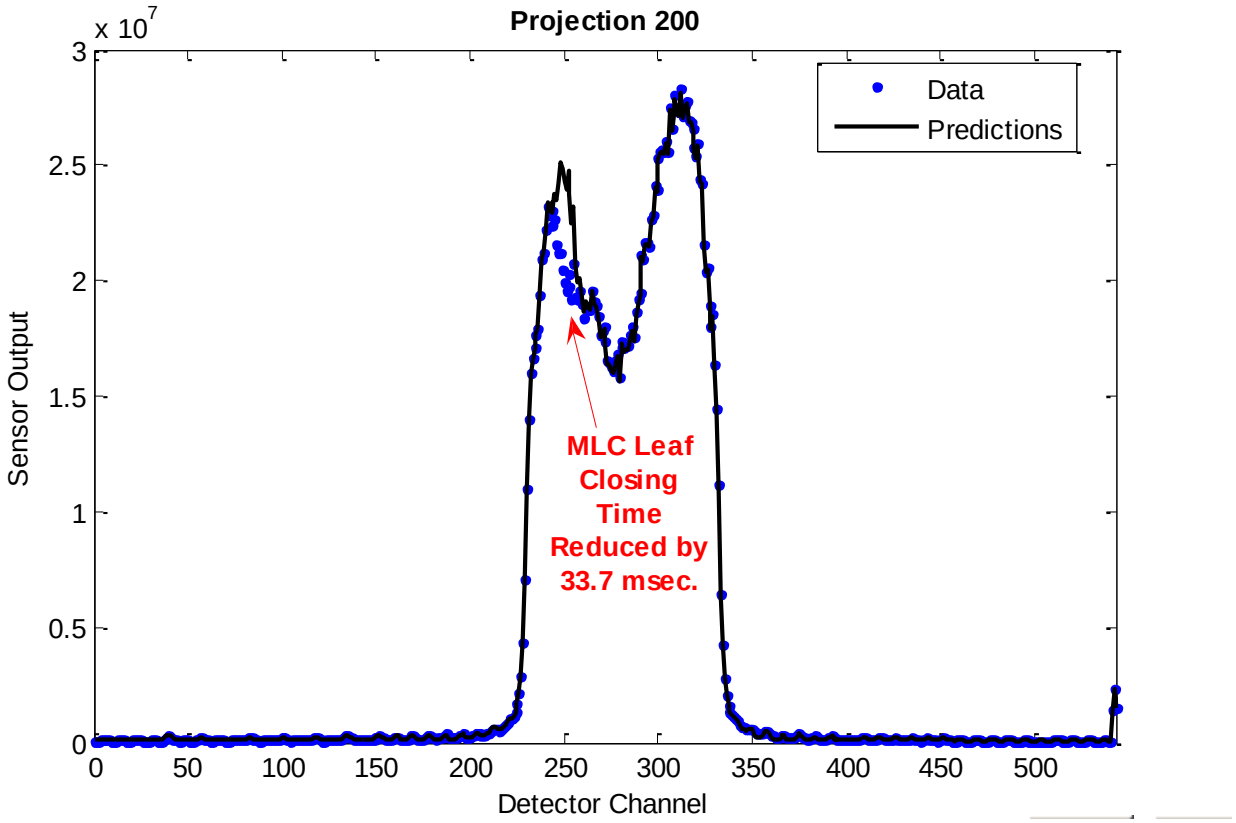

Projection 298
x 10^7
Sensor Output
Detector Channel
Data
Predictions
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
0
50
100
150
200
250
300
350
400
450
500

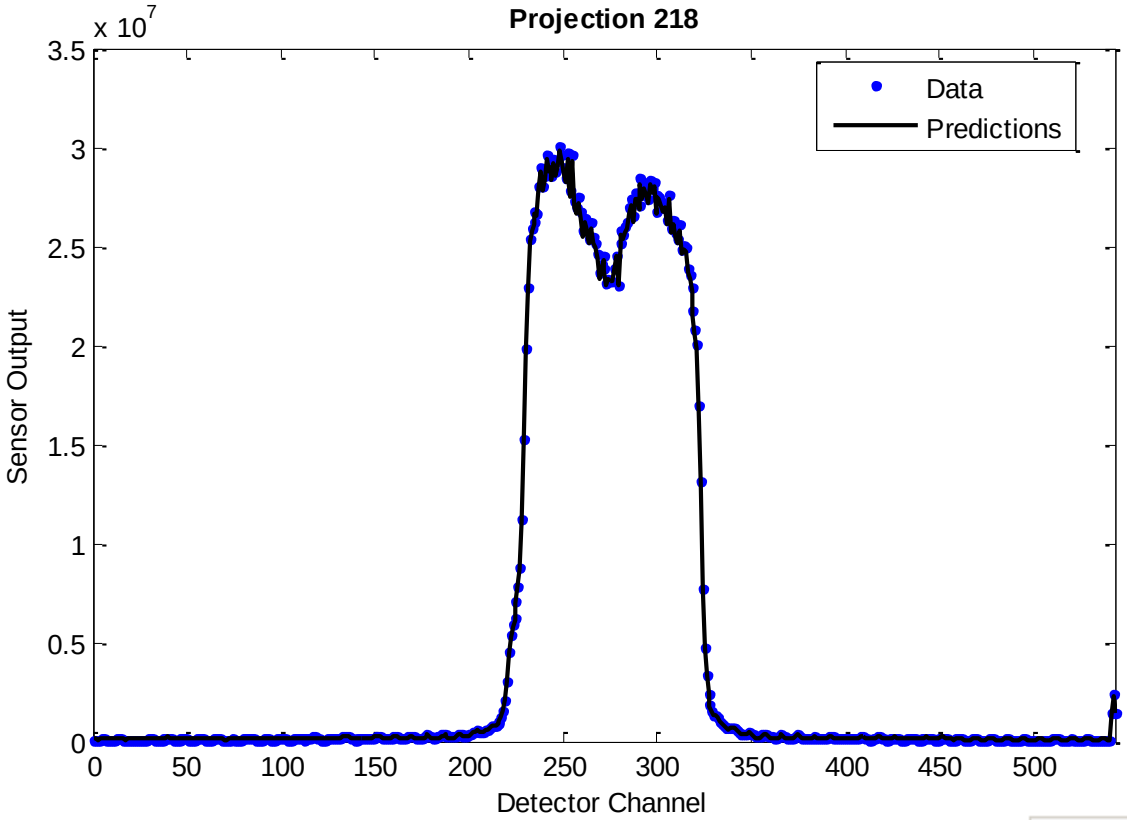

Projection 299
x 10^7
Sensor Output
Detector Channel
Data
Predictions
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
0
50
100
150
200
250
300
350
400
450
500

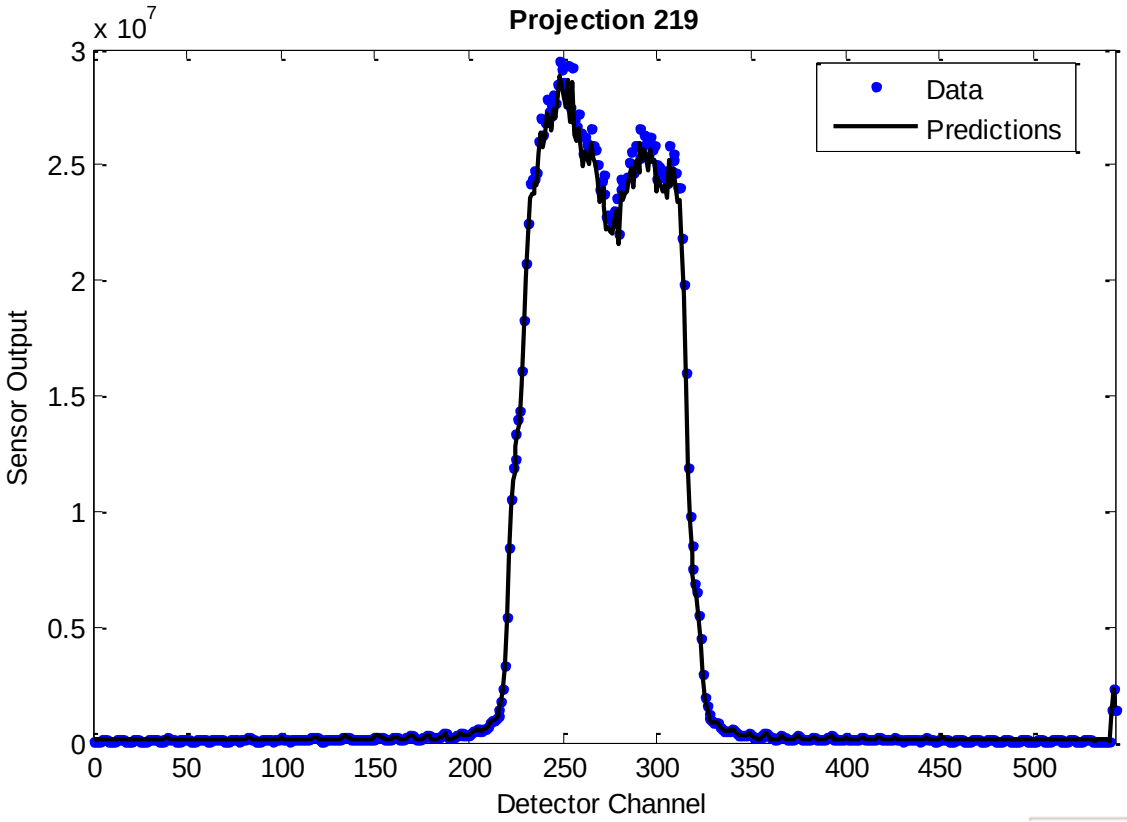

Projection 300
x 10^7
Sensor Output
Detector Channel
Data
Predictions
MLC Leaf Closing Time Reduction by 20.8 msec

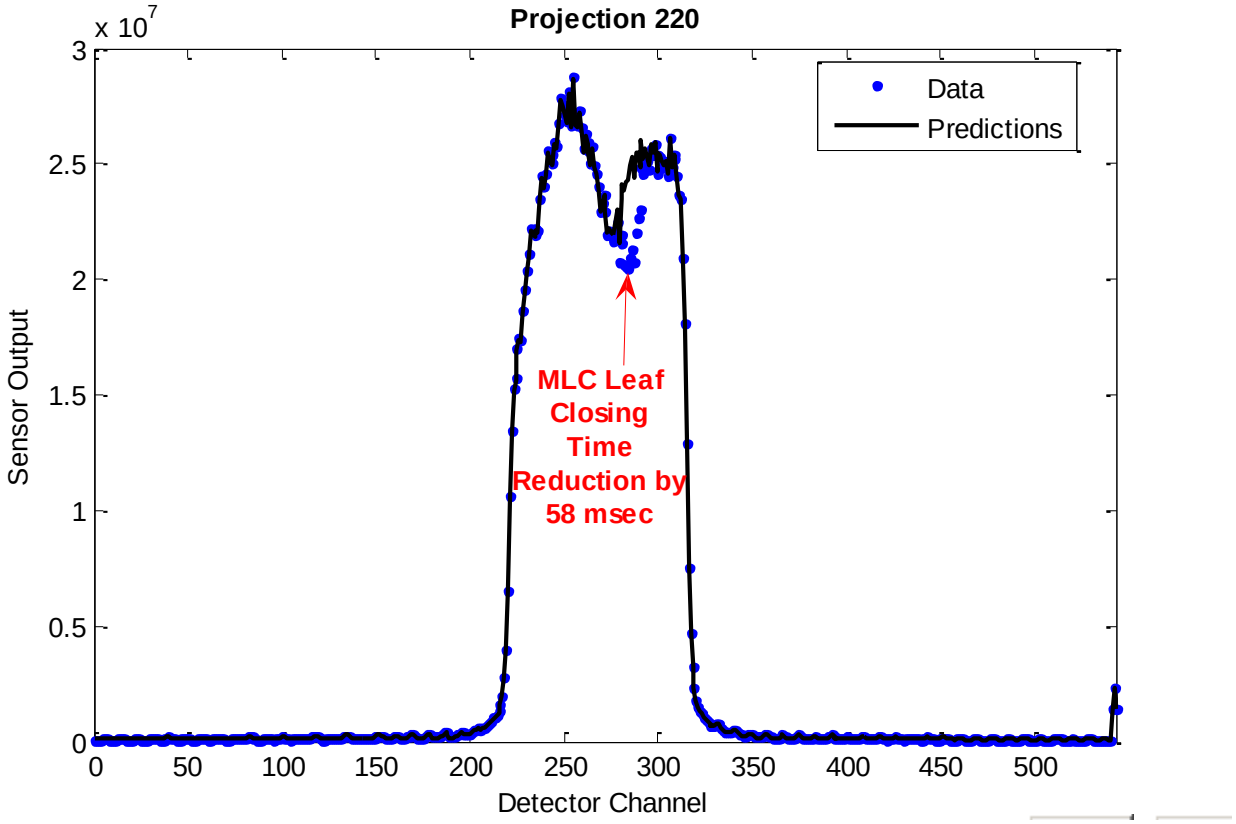

Projection 198
x 10^7
Sensor Output
Detector Channel
Data
Predictions

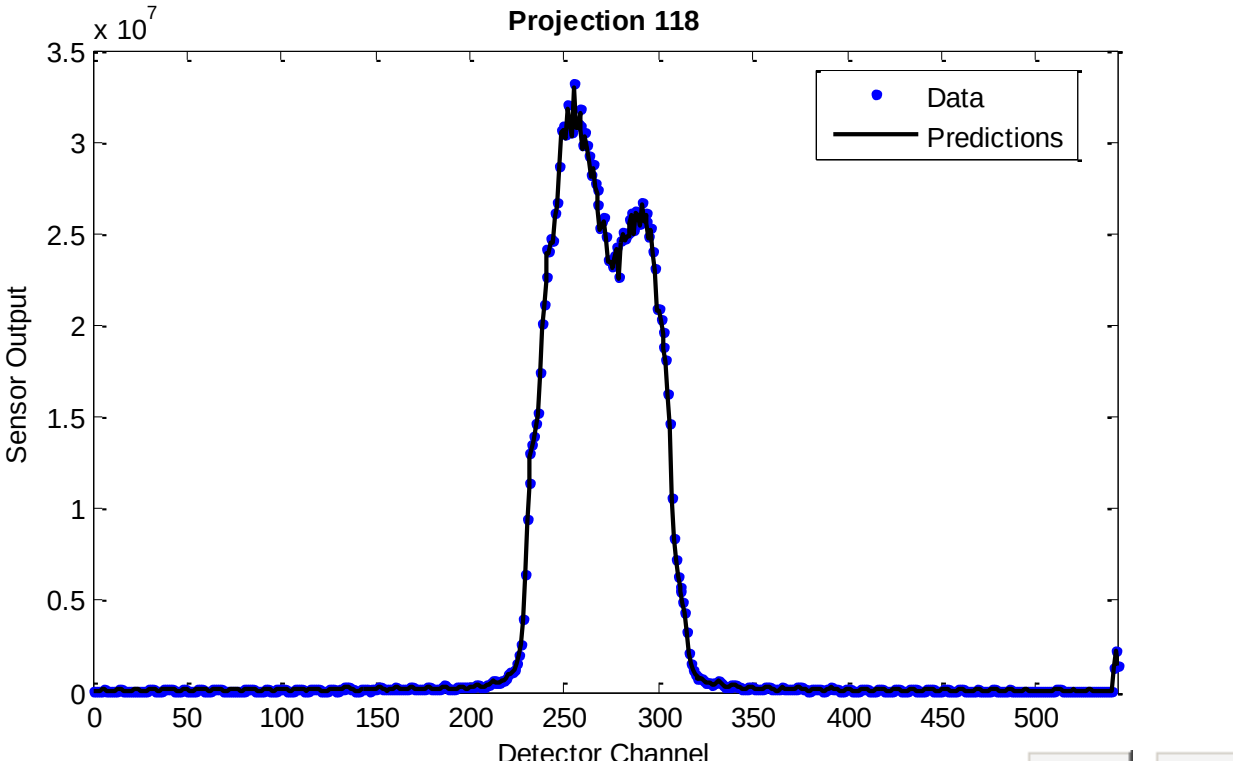

Projection 199
x 10^7
Sensor Output
Detector Channel
Data
Predictions
0
0.5
1
1.5
2
2.5
3
0
50
100
150
200
250
300
350
400
450
500

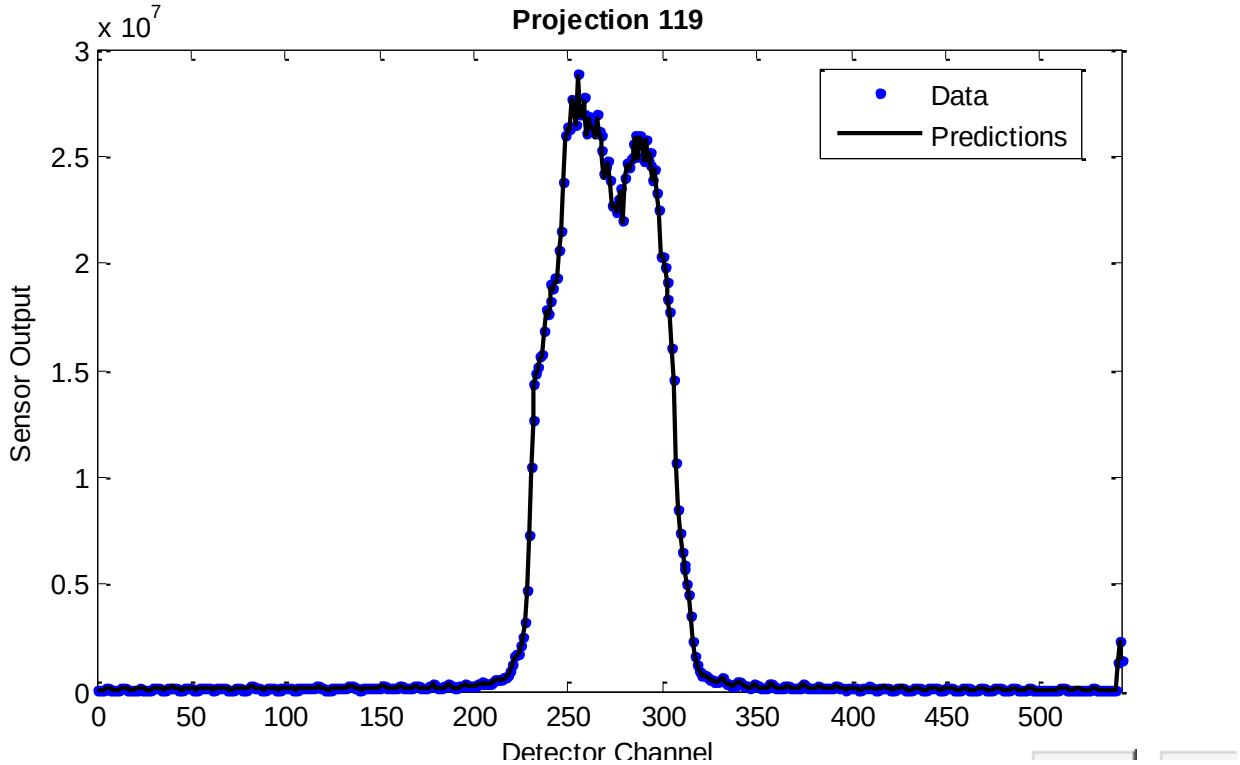

Projection 200
x 10^7
Sensor Output
Detector Channel
Data
Predictions
MLC Leaf Closing Time Reduced by 33.7 msec.
0
0.5
1
1.5
2
2.5
3
0
50
100
150
200
250
300
350
400
450
500

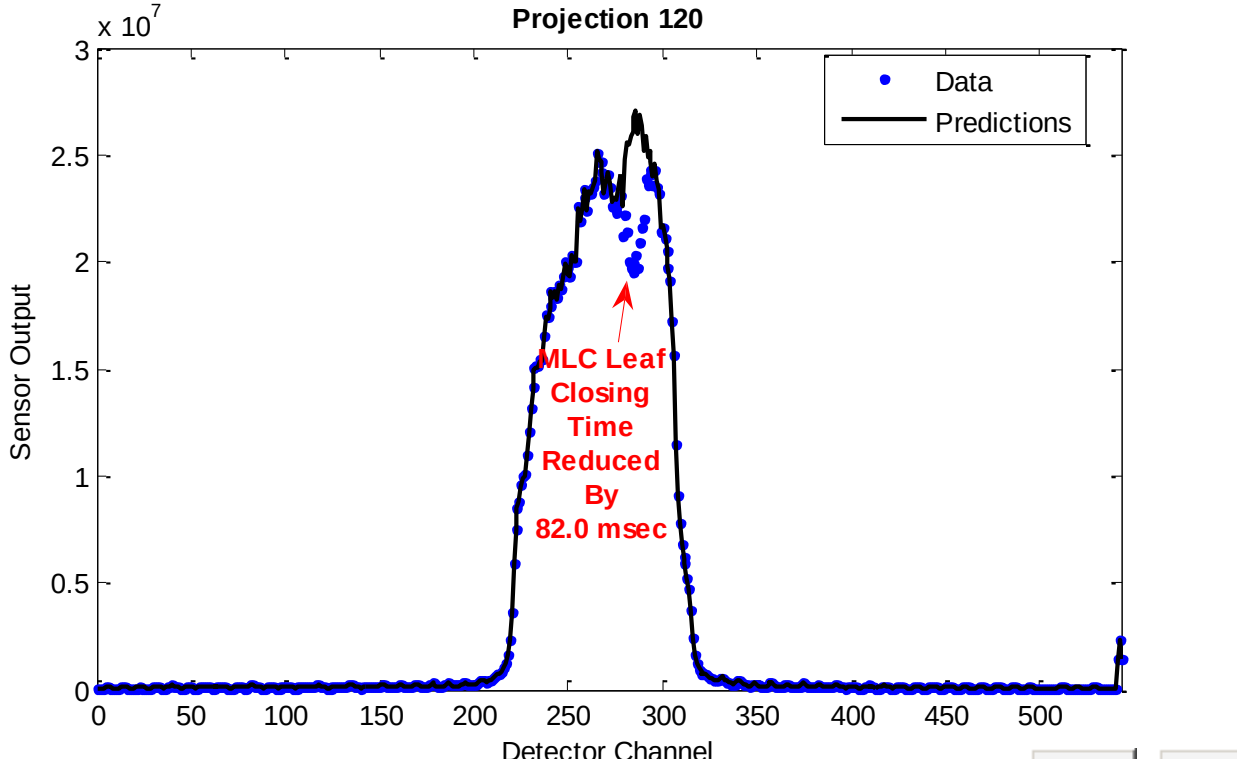

Projection 218
x 10^7
Sensor Output
Detector Channel
Data
Predictions
0
0.5
1
1.5
2
2.5
3
3.5
0
50
100
150
200
250
300
350
400
450
500


Projection 219
x 10^7
Sensor Output
Detector Channel
Data
Predictions
0
0.5
1
1.5
2
2.5
3
0
50
100
150
200
250
300
350
400
450
500


Projection 220
x 10^7
Sensor Output
Detector Channel
Data
Predictions
MLC Leaf Closing Time Reduction by 58 msec
0
0.5
1
1.5
2
2.5
3
0
50
100
150
200
250
300
350
400
450
500


Projection 118
x 10^7
Sensor Output
Detector Channel
Data
Predictions
0
0.5
1
1.5
2
2.5
3
3.5
0
50
100
150
200
250
300
350
400
450
500


Projection 119
x 10^7
Sensor Output
Detector Channel
Data
Predictions
0
0.5
1
1.5
2
2.5
3
0
50
100
150
200
250
300
350
400
450
500


Projection 120
x 10^7
Sensor Output
Detector Channel
Data
Predictions
MLC Leaf Closing Time Reduced By 82.0 msec
0
0.5
1
1.5
2
2.5
3
0
50
100
150
200
250
300
350
400
450
500

VITA

Rebecca Marie Seibert was born and raised in Pinckneyville, IL. For high school, she attended the Illinois Mathematics and Science Academy in Aurora IL. After graduating high school, she attended the Georgia Institute of Technology. At Georgia Tech, she was involved in many extracurricular activities including SWE, Alpha Chi Omega, Alpha Chi Sigma, tutoring at Centennial Place Elementary, and participated in two study abroad programs, travelling to Germany, and then Australia, New Zealand, and Japan. While at Georgia Tech, Rebecca met her future husband Ryan Graciano, whom she would later marry in August of 2008. She graduated from Georgia Tech in 2004 with a Bachelor of Science in Chemical Engineering.

She then attended graduate school at the University of Tennessee, Knoxville. She received her Masters of Science in Nuclear Engineering in May, 2006 under the guidance of Dr. J Wesley Hines, Ph.D.. While completing her Masters, she worked as a graduate research assistant and also volunteered at the Thompson Cancer Survival Center as a Medical Physics researcher. After a year of volunteering, she was hired at the cancer center as a medical physics resident under the supervision of Dr. Chester Ramsey, Ph.D.. Throughout her residency, she continued to pursue a PhD by conducting research and taking evening classes. With the support and encouragement of Dr. Hines and Dr. Ramsey, she was able to apply novel modeling, fault detection, and optimization techniques to the field of radiation therapy. After finishing her residency and completing all of the necessary coursework, she took a medical physics position with West County Radiology in St. Louis, MO to be closer to her family. While in St. Louis, she passed all of the board exams to become a diplomat of the American Board of Radiology. In August 2012, she and her husband moved to California. She is now working as a medical

physicist at the John Muir Cancer Center in Walnut Creek, CA. She will receive a Doctor of Philosophy degree in Nuclear Engineering from the University of Tennessee in December 2012. She hopes to continue her research and find new ways that nonparametric modeling techniques can be used to improve radiation therapy.