

5-2014

Forecasting Longitudinal Changes in Oropharyngeal Tumor Volume, Position, and Morphology during Image-Guided Radiation Therapy

Adam D. Yock

Follow this and additional works at: https://digitalcommons.library.tmc.edu/utgsbs_dissertations



Part of the [Medicine and Health Sciences Commons](#), [Other Physical Sciences and Mathematics Commons](#), and the [Other Physics Commons](#)

Recommended Citation

Yock, Adam D., "Forecasting Longitudinal Changes in Oropharyngeal Tumor Volume, Position, and Morphology during Image-Guided Radiation Therapy" (2014). *The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access)*. 438.

https://digitalcommons.library.tmc.edu/utgsbs_dissertations/438

This Dissertation (PhD) is brought to you for free and open access by the The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences at DigitalCommons@TMC. It has been accepted for inclusion in The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access) by an authorized administrator of DigitalCommons@TMC. For more information, please contact digitalcommons@library.tmc.edu.

FORECASTING LONGITUDINAL CHANGES
IN OROPHARYNGEAL TUMOR VOLUME, POSITION, AND MORPHOLOGY
DURING IMAGE-GUIDED RADIATION THERAPY

by

Adam Douglas Yock, A.B.

APPROVED:

Laurence E. Court, Ph.D.
Supervisory Professor

Beth M. Beadle, M.D., Ph.D.

Lei Dong, Ph.D.

Rajat J. Kudchadker, Ph.D.

Arvind Rao, Ph.D.

APPROVED:

Dean, The University of Texas
Graduate School of Biomedical Sciences at Houston

FORECASTING LONGITUDINAL CHANGES
IN OROPHARYNGEAL TUMOR VOLUME, POSITION, AND MORPHOLOGY
DURING IMAGE-GUIDED RADIATION THERAPY

A
DISSERTATION

Presented to the Faculty of
The University of Texas
Health Science Center at Houston
and
The University of Texas
MD Anderson Cancer Center
Graduate School of Biomedical Sciences
in Partial Fulfillment
of the Requirements
for the degree of

DOCTOR OF PHILOSOPHY

by

Adam Douglas Yock, A.B.
Houston Texas

May, 2014

Dedicated with love to Douglas, Doris, and Lindsey Yock,
who have supported me patiently for years as I've tried to fit square pegs into round holes.

ACKNOWLEDGEMENTS

I first want to thank my advisor, Dr. Laurence Court, for his unwarranted faith in me and in my work. His enthusiasm and patience amidst high-dimensional abstractness facilitated the success of this project. His perspective on research and mentorship is inspirational and resonates strongly with his students.

I would also like to thank Dr. Arvind Rao for his persistent energy and willingness to help. His knowledge of image analysis was fundamental to this work.

I would like to thank Dr. Lei Dong for taking me on as a student and providing a rich idea-space in which to conceive a meaningful project. His experience and advice regarding research and graduate school were invaluable.

I would like to thank Dr. Beth Beadle and Dr. Rajat Kudchadker for their guidance in steering my dissertation research and for their efforts on numerous abstracts and manuscripts.

I would like to thank Dr. Jinzhong Yang and Dr. Lifei Zhang for their innumerable contributions and suggestions regarding deformable image registration and image processing and analysis. Us students would flounder without their support.

I would like to thank my classmates and group-mates for making graduate school vibrant and enjoyable. I would especially like to thank Joey Cheung for literally being by my side for the past several years and listening to a constant barrage of ridiculous ideas and questions lobbed in his direction.

Lastly, I would like to thank my parents and sister. Their unwavering love, patience, and support are behind every step I take in the right direction.

ABSTRACT

The purpose of this work was to generate, evaluate, and compare models that predict longitudinal changes in oropharyngeal tumor volume, position, and morphology during radiation therapy.

One volume, one position, and two morphology (size, shape, and position) feature vectors were used to describe 35 oropharyngeal gross tumor volumes (GTVs) during radiation therapy. The two morphology feature vectors comprised the coordinates of the GTV centroids and one of two shape descriptors. One shape descriptor was based on radial distances between the GTV centroid and 614 surface landmarks. The other was based on a spherical harmonic decomposition of these distances. For a training set of patients, the changes in feature vectors during treatment were represented by static, linear, mean, and median models along with two models derived from principal component analysis. The error of these models in forecasting the GTV volume, position, and morphology of a test patient was evaluated using leave-one-out cross-validation, and the accuracy of the models were compared with Wilcoxon signed-rank tests. The effect on accuracy of adjusting model parameters at 1, 2, 3, or 5 time points (“adjustment points”) was also evaluated.

Including a single adjustment point improved the accuracy in forecasting the volume and position by 12.9% – 30.0% and 27.0% – 34.1%, respectively. For the two morphology feature vectors, a single adjustment point improved accuracy by 27.5 – 33.8% and 28.6% – 33.0%. Additional adjustment points further improved accuracy, but with diminishing effects. Non-static models demonstrated greater accuracy than static models with equal numbers of adjustment points. These improvements were small, except for models of the volume (54.2% – 58.0%) for which they were greater than those of adding an adjustment point. For the other three feature vectors, the effect of including an adjustment point was greater than that of selecting a non-static model.

Tumor volume, position, and morphology were predicted at each treatment fraction using models that include information from prior patients and/or prior treatment fractions. The predicted tumor morphology can be compared with patient anatomy or dose distributions, thereby providing a more complete depiction of treatment response, influencing clinical decision making, and opening the possibility of anticipatory re-planning.

TABLE OF CONTENTS

Dedication	iv
Acknowledgements	v
Abstract	vi
Table of Contents	vii
List of Illustrations	x
List of Tables	xix
Chapter 1 – Introduction	1
Conventional predictive model limitation (i).....	4
Conventional predictive model limitation (ii).....	4
Chapter 2 – Principal Hypothesis and Specific Aims	8
Principal Hypothesis	8
Specific Aim 1.....	8
Specific Aim 2	8
Specific Aim 3	9
Specific Aim 4	9
Chapter 3 – Methodology	10
3.1. Instances of the gross tumor volume and deformable image registration	10
3.2. Descriptor variables	13
3.2.1. GTV volume descriptor	13
3.2.2. Bony anatomy landmarks	14
3.2.3. GTV centroid position descriptor	17
3.2.4. Radial extent GTV morphology descriptor.....	17

3.2.5. Spherical harmonic GTV morphology descriptor.....	18
3.2.6. Regularizing feature vector elements.....	20
3.3. Forecast models	21
3.3.1. Static model	21
3.3.2. Linear model	21
3.3.3. Mean model	22
3.3.4. Median model	22
3.3.5. Vector PCA and 2D PCA models.....	23
3.4. Representing changes in feature vector elements	25
3.5. Model adjustment schemes	25
3.6. Evaluation of model forecast accuracy with leave-one-out cross-validation	28
3.7. Statistical comparison of model forecast accuracy	29
Chapter 4 – Results	31
4.1. Generating feature vectors	31
4.2. Results for Specific Aims	35
4.2.1. Specific Aim 1 results.....	46
4.2.1.1. Specific Aim 1 results – GTV volume descriptor.....	46
4.2.1.2. Specific Aim 1 results – GTV centroid position descriptor....	54
4.2.2. Specific Aim 2 results.....	62
4.2.2.1. Specific Aim 2 results – Radial extent GTV morphology descriptor.....	62
4.2.2.2. Specific Aim 2 results – Spherical harmonic GTV morphology descriptor.....	70

4.2.3. Specific Aim 3 results	78
4.2.3.1. Specific Aim 3 results – GTV volume descriptor.....	78
4.2.3.2. Specific Aim 3 results – GTV centroid position descriptor....	82
4.2.4. Specific Aim 4 results	86
4.2.4.1. Specific Aim 4 results – Radial extent GTV morphology descriptor.....	86
4.2.4.2. Specific Aim 4 results – Spherical harmonic GTV morphology descriptor.....	90
4.3. Reconstructed forecast morphology	94
Chapter 5 – Discussion	97
Chapter 6 – Conclusion.....	108
Bibliography	109
Vita.....	121

LIST OF ILLUSTRATIONS

Figure 1: CT-on-rails system	11
Figure 2: Locations of bony anatomy landmarks	15
Figure 3: Spherical harmonics.	19
Figure 4: Diagram of model adjustment schemes.....	26
Figure 5: Illustration of static and linear models with and without adjustment points.....	28
Figure 6: Morphology approximation error according to the maximum degree of the spherical harmonic basis	32
Figure 7: Smoothing spline error – GTV volume descriptor	33
Figure 8: Smoothing spline error – GTV centroid position descriptor	33
Figure 9: Smoothing spline error – Radial extent GTV morphology descriptor	34
Figure 10: Forecast error of static and linear models – GTV volume descriptor	35
Figure 11: Forecast error of static and mean models – GTV volume descriptor.....	36
Figure 12: Forecast error of static and median models – GTV volume descriptor.....	36
Figure 13: Forecast error of static and vector PCA models – GTV volume descriptor.....	37
Figure 14: Forecast error of static and 2D PCA models – GTV volume descriptor.....	37
Figure 15: Forecast error of static and linear models – GTV centroid position descriptor	38
Figure 16: Forecast error of static and mean models – GTV centroid position descriptor.....	38
Figure 17: Forecast error of static and median models – GTV centroid position descriptor....	39
Figure 18: Forecast error of static and vector PCA models – GTV centroid position descriptor	39
Figure 19: Forecast error of static and 2D PCA models – GTV centroid position descriptor	40

Figure 20: Forecast error of static and linear models – Radial extent GTV morphology	
descriptor	40
Figure 21: Forecast error of static and mean models – Radial extent GTV morphology	
descriptor	41
Figure 22: Forecast error of static and median models – Radial extent GTV morphology	
descriptor	41
Figure 23: Forecast error of static and vector PCA models – Radial extent GTV morphology	
descriptor	42
Figure 24: Forecast error of static and 2D PCA models – Radial extent GTV morphology	
descriptor	42
Figure 25: Forecast error of static and linear models – Spherical harmonic GTV morphology	
descriptor	43
Figure 26: Forecast error of static and mean models – Spherical harmonic GTV morphology	
descriptor	43
Figure 27: Forecast error of static and median models – Spherical harmonic GTV	
morphology descriptor	44
Figure 28: Forecast error of static and vector PCA models – Spherical harmonic GTV	
morphology descriptor	44
Figure 29: Forecast error of static and 2D PCA models – Spherical harmonic GTV	
morphology descriptor	45
Figure 30: Change in forecast error of the static model with adjustment points – GTV volume	
descriptor	46

Figure 31: Change in forecast error of the linear model with adjustment points – GTV	
volume descriptor	47
Figure 32: Change in forecast error of the mean model with adjustment points – GTV volume	
descriptor	47
Figure 33: Change in forecast error of the median model with adjustment points – GTV	
volume descriptor	48
Figure 34: Change in forecast error of the vector PCA model with adjustment points – GTV	
volume descriptor	48
Figure 35: Change in forecast error of the 2D PCA model with adjustment points – GTV	
volume descriptor	49
Figure 36: Marginal change in forecast error of the static model with subsequent adjustment	
points – GTV volume descriptor	50
Figure 37: Marginal change in forecast error of the linear model with subsequent adjustment	
points – GTV volume descriptor	51
Figure 38: Marginal change in forecast error of the mean model with subsequent adjustment	
points – GTV volume descriptor	51
Figure 39: Marginal change in forecast error of the median model with subsequent	
adjustment points – GTV volume descriptor	52
Figure 40: Marginal change in forecast error of the vector PCA model with subsequent	
adjustment points – GTV volume descriptor	52
Figure 41: Marginal change in forecast error of the 2D PCA model with subsequent	
adjustment points – GTV volume descriptor	53

Figure 42: Change in forecast error of the static model with adjustment points – GTV

centroid position descriptor54

Figure 43: Change in forecast error of the linear model with adjustment points – GTV

centroid position descriptor55

Figure 44: Change in forecast error of the mean model with adjustment points – GTV

centroid position descriptor55

Figure 45: Change in forecast error of the median model with adjustment points – GTV

centroid position descriptor56

Figure 46: Change in forecast error of the vector PCA model with adjustment points – GTV

centroid position descriptor56

Figure 47: Change in forecast error of the 2D PCA model with adjustment points – GTV

centroid position descriptor57

Figure 48: Marginal change in forecast error of the static model with subsequent adjustment

points – GTV centroid position descriptor58

Figure 49: Marginal change in forecast error of the linear model with subsequent adjustment

points – GTV centroid position descriptor59

Figure 50: Marginal change in forecast error of the mean model with subsequent adjustment

points – GTV centroid position descriptor59

Figure 51: Marginal change in forecast error of the median model with subsequent

adjustment points – GTV centroid position descriptor60

Figure 52: Marginal change in forecast error of the vector PCA model with subsequent

adjustment points – GTV centroid position descriptor60

Figure 53: Marginal change in forecast error of the 2D PCA model with subsequent adjustment points – GTV centroid position descriptor	61
Figure 54: Change in forecast error of the static model with adjustment points – Radial extent GTV morphology descriptor	62
Figure 55: Change in forecast error of the linear model with adjustment points – Radial extent GTV morphology descriptor	63
Figure 56: Change in forecast error of the mean model with adjustment points – Radial extent GTV morphology descriptor	63
Figure 57: Change in forecast error of the median model with adjustment points – Radial extent GTV morphology descriptor	64
Figure 58: Change in forecast error of the vector PCA model with adjustment points – Radial extent GTV morphology descriptor	64
Figure 59: Change in forecast error of the 2D PCA model with adjustment points – Radial extent GTV morphology descriptor	65
Figure 60: Marginal change in forecast error of the static model with subsequent adjustment points – Radial extent GTV morphology descriptor	66
Figure 61: Marginal change in forecast error of the linear model with subsequent adjustment points – Radial extent GTV morphology descriptor	67
Figure 62: Marginal change in forecast error of the mean model with subsequent adjustment points – Radial extent GTV morphology descriptor	67
Figure 63: Marginal change in forecast error of the median model with subsequent adjustment points – Radial extent GTV morphology descriptor	68

Figure 64: Marginal change in forecast error of the vector PCA model with subsequent adjustment points – Radial extent GTV morphology descriptor	68
Figure 65: Marginal change in forecast error of the 2D PCA model with subsequent adjustment points – Radial extent GTV morphology descriptor	69
Figure 66: Change in forecast error of the static model with adjustment points – Spherical harmonic GTV morphology descriptor	70
Figure 67: Change in forecast error of the linear model with adjustment points – Spherical harmonic GTV morphology descriptor	71
Figure 68: Change in forecast error of the mean model with adjustment points – Spherical harmonic GTV morphology descriptor	71
Figure 69: Change in forecast error of the median model with adjustment points – Spherical harmonic GTV morphology descriptor	72
Figure 70: Change in forecast error of the vector PCA model with adjustment points – Spherical harmonic GTV morphology descriptor	72
Figure 71: Change in forecast error of the 2D PCA model with adjustment points – Spherical harmonic GTV morphology descriptor	73
Figure 72: Marginal change in forecast error of the static model with subsequent adjustment points – Spherical harmonic GTV morphology descriptor	74
Figure 73: Marginal change in forecast error of the linear model with subsequent adjustment points – Spherical harmonic GTV morphology descriptor	75
Figure 74: Marginal change in forecast error of the mean model with subsequent adjustment points – Spherical harmonic GTV morphology descriptor	75

Figure 75: Marginal change in forecast error of the median model with subsequent adjustment points – Spherical harmonic GTV morphology descriptor.....	76
Figure 76: Marginal change in forecast error of the vector PCA model with subsequent adjustment points – Spherical harmonic GTV morphology descriptor.....	76
Figure 77: Marginal change in forecast error of the 2D PCA model with subsequent adjustment points – Spherical harmonic GTV morphology descriptor.....	77
Figure 78: Change in forecast error from the static model to the linear model with an equal number of adjustment points – GTV volume descriptor	79
Figure 79: Change in forecast error from the static model to the mean model with an equal number of adjustment points – GTV volume descriptor	79
Figure 80: Change in forecast error from the static model to the median model with an equal number of adjustment points – GTV volume descriptor	80
Figure 81: Change in forecast error from the static model to the vector PCA model with an equal number of adjustment points – GTV volume descriptor	80
Figure 82: Change in forecast error from the static model to the 2D PCA model with an equal number of adjustment points – GTV volume descriptor	81
Figure 83: Change in forecast error from the static model to the linear model with an equal number of adjustment points – GTV centroid position descriptor	83
Figure 84: Change in forecast error from the static model to the mean model with an equal number of adjustment points – GTV centroid position descriptor	83
Figure 85: Change in forecast error from the static model to the median model with an equal number of adjustment points – GTV centroid position descriptor	84

Figure 86: Change in forecast error from the static model to the vector PCA model with an equal number of adjustment points – GTV centroid position descriptor	84
Figure 87: Change in forecast error from the static model to the 2D PCA model with an equal number of adjustment points – GTV centroid position descriptor	85
Figure 88: Change in forecast error from the static model to the linear model with an equal number of adjustment points – Radial extent GTV morphology descriptor	87
Figure 89: Change in forecast error from the static model to the mean model with an equal number of adjustment points – Radial extent GTV morphology descriptor	87
Figure 90: Change in forecast error from the static model to the median model with an equal number of adjustment points – Radial extent GTV morphology descriptor	88
Figure 91: Change in forecast error from the static model to the vector PCA model with an equal number of adjustment points – Radial extent GTV morphology descriptor	88
Figure 92: Change in forecast error from the static model to the 2D PCA model with an equal number of adjustment points – Radial extent GTV morphology descriptor	89
Figure 93: Change in forecast error from the static model to the linear model with an equal number of adjustment points – Spherical harmonic GTV morphology descriptor	91
Figure 94: Change in forecast error from the static model to the mean model with an equal number of adjustment points – Spherical harmonic GTV morphology descriptor	91
Figure 95: Change in forecast error from the static model to the median model with an equal number of adjustment points – Spherical harmonic GTV morphology descriptor	92
Figure 96: Change in forecast error from the static model to the vector PCA model with an equal number of adjustment points – Spherical harmonic GTV morphology descriptor	92

Figure 97: Change in forecast error from the static model to the 2D PCA model with an equal number of adjustment points – Spherical harmonic GTV morphology descriptor	93
Figure 98: Forecast contours of static models with different numbers of adjustment points	95
Figure 99: Forecast contours of static and linear models with no adjustment points	96

LIST OF TABLES

Table I: Illustration of the relationship between the four Specific Aims.....	9
Table II: Patient information.....	11
Table III: Summary of the results for the four Specific Aims	103

“Prediction is very difficult, especially about the future.”

— Niels Bohr

Chapter 1

INTRODUCTION

Cancer of the head and neck comprises malignant disease of the nasal cavity, oral cavity, pharynx, larynx, and sinuses. There were 53,640 new cases of head and neck cancer (3.2% of all new cancer cases) along with 11,520 deaths (2.0% of all cancer deaths) estimated for 2013 in the US.¹ These malignancies are predominantly squamous cell carcinoma arising in the epithelium of the upper aerodigestive tract.² Primary risk factors include alcohol and tobacco (implicated in 75% of cases) which also exhibit a synergistic effect.³⁻⁵ The role of human papillomavirus (HPV) – particularly type 16 and, to an extent, type 18 – as an additional risk factor has become clear, implicated in 25% to 35% of cases.^{6, 7}

A quarter of the new head and neck cancer cases and a fifth of the associated deaths are attributable to disease of the oropharynx, primarily of the tonsils and the base of tongue.¹ Two-thirds of these present at an advanced stage, though relatively few present with distant metastasis.¹ Unlike disease staging for other head and neck cancers which depends on the extent of disease involvement, staging for cancers of the oropharynx is largely based on the physical dimensions of the observed tumor. Five-year survival rates for localized, regional, and distant malignant disease of the oropharynx are 82.7%, 59.2%, and 36.3%, respectively.¹

Surgery and radiation remain the principal modalities for treating oropharyngeal cancer, though systemic therapies concurrent with radiation provide an additional benefit.⁸⁻¹⁵

Radiation therapy is generally delivered with external, megavoltage, x-ray beams.

Commonly, approximately 2 Gy fractional doses are delivered for several weeks to achieve a total dose of 60 Gy to 70 Gy depending on disease stage.⁶ Hyperfractionated and accelerated fractionation schemes have also been explored.^{6, 10, 11, 13-15} In addition, interstitial radioactive

implants have been used to provide a brachytherapy boost for tumors, primarily those of the base of tongue.⁶

Treatment plans for fractionated radiation therapy are typically generated based on the patient anatomy as depicted in a single, pre-treatment computed tomography (CT) image set. Linear accelerator beam parameters that produce the optimal dose distribution on this planning CT are used to administer the radiation at each treatment fraction thereafter. However, anatomic changes and motion, along with limited patient positioning reproducibility, cause the radiation to be administered to anatomy dissimilar to that of the planning CT. These anatomic variations affect the dose distribution delivered to the patient, potentially compromising the therapeutic benefit and clinical safety of the treatment.

The uncertainty regarding variations in patient anatomy and positioning can manifest in multiple forms. Instances of flexible bony anatomy and deformable soft tissue anatomy may often be considered random. Similarly, the precision with which a patient is positioned for each treatment fraction is subject to random error. Governed by chance, the effect of these random variations between each treatment fraction is to blur the intended dose distribution, flattening steep dose gradients characteristic of highly-conformal, intensity-modulated radiation therapy treatment plans.^{16, 17}

The planning CT is also subject to random variations. In this case however, the particular anatomic configuration is used to establish a baseline. Any discrepancy between the value of a random variable observed during acquisition of the planning CT and its average value during treatment perpetuates throughout treatment as a systematic variation. Unlike the blurring effect of random variations occurring between fractions, systematic variations shift the delivered dose distribution relative to the intended one. This shift is particularly

consequential in regions with steep dose gradients where even a small shift could lead to a large change in the delivered dose.

Other anatomic variations are correlated among treatment fractions and are thus inherently non-random. These include longitudinal trends in anatomic changes that progress during treatment, such as changes in the volume and position of the tumor and other organs.¹⁸ Longitudinal changes may be due directly to the treatment, or in conjunction with other changes like weight loss or the resolution of edema or other post-surgical effects.

The clinical implications of delivering an unintended radiation dose distribution have led to extensive research to quantify the uncertainties associated with these variations. Numerous uncertainty estimates have been derived using nearly as many methodologies. Nonetheless, random and systematic variations between treatment fractions are typically 1 mm to 3 mm in magnitude, while those occurring within a single fraction are less than 2 mm.¹⁹⁻²⁸ Castadot reviewed literature describing the dosimetric consequences of these uncertainties, such as a decrease in target dose and an increase in normal tissue dose.²⁹

In order to maintain the intended dose distribution, considerable effort is made to reduce the magnitude of these variations. Strategies include immobilizing the patient, adding geometric margins around tumors during treatment planning, imaging the patient prior to each treatment fraction, and adjusting the radiation therapy plan during the course of treatment. The latter strategy is referred to as adaptive radiation therapy (ART). It involves re-imaging the patient and re-optimizing the treatment plan in order to tailor a new plan to the longitudinal changes in patient anatomy.

Developed by Yan et al.,³⁰ the workflow and clinical application of ART have been described and assessed by many others.³¹⁻³⁶ However, practical considerations have restricted

the clinical utility of ART. In principle, an image of the patient acquired before each treatment fraction could be used to generate a new treatment plan to be administered exclusively on that fraction. This workflow would promote confidence in the fidelity of the delivered dose distribution to the intended one. However, re-imaging and re-optimizing frequently during treatment currently remains prohibitively time and computationally intensive. As a result, most implementations of ART are limited to one or a few treatment plan adjustments. Another limitation, closely related to the limited number of adaptive adjustments, is that adjustments are inherently reactionary. Unless achieved in real time, these interventions necessarily lag behind their impetus. When only a small number of treatment plan adjustments are made, the time discrepancy between the re-optimization of a new plan and its administration can grow to days or weeks. Consequences of this lag may be diminished by anticipating anatomic changes. A predictive model of longitudinal changes in anatomy could lead to treatment plan adjustments and interventions less subject to the effects of these trends.

Anticipating changes in descriptive anatomic variables has been considered previously. Predictions have been made regarding tumor volume, position, and surface area.³⁷⁻⁴⁰ In addition, contour slice area and curvature,^{41, 42} central axis position,⁴³ and mathematical notions of sphericity⁴⁴ and eccentricity⁴⁵ have also been used to describe radiation therapy targets and organs at risk. However, the application of conventional predictive models is usually restricted by at least one of the following two limitations:

- (i) Predictive models often only describe the overall change in the descriptor variable.
- (ii) Descriptor variables are typically reductive and non-generative.

The first of these limitations treats the response of the patient as a state function specified according to its initial and final state. This concept ignores clinical insight that might be garnered from the manner in which the anatomy changes, beyond that available from the total change. A more complete description is to consider longitudinal anatomic changes as a time series representing a continuous function. This functional description can resolve differences in the response of two patients who demonstrate similar total changes, but have responded at different times and at different rates. Nuances in treatment response could reflect differences in disease progression, histology, microenvironment, or resistance. The manner in which the anatomy changes also affects the radiation dose to nearby normal tissues and warrants varied implementation of ART.

The second of the two limitations restricts the utility of predictive models. Reductive (or non-generative) descriptors are those that irreversibly discard information, typically for the sake of simplicity and/or efficiency. The lost information, however, renders it impossible to specifically reconstruct the original object from the descriptor. This limits the application of these models compared to models based on non-reductive, generative descriptors.

Challenges associated with retaining sufficient information for a descriptor to be generative include higher variable dimensionality and greater analytic complexity. As a result, successful analysis requires increased data sample sizes and computational resources. Nonetheless, uses for generative descriptors abound. A number of them are described in reviews by Iyer et al.⁴⁶ and Tangelder et al.⁴⁷

When applied to anatomy, generative models retain information regarding the *morphology* (size, shape, and position) of anatomic features. Applications of such descriptors in medicine include neurosurgery⁴⁸ and computational neuroanatomy,⁴⁹ along with computer-

aided diagnosis of breast cancer.^{50, 51} In radiation therapy, a number of authors have used various reductive and generative descriptors as shape models representing random anatomic variations. Generative descriptors in radiation therapy included biomechanical finite element models,^{38, 52, 53} deformable image registration derived vector models,⁵⁴ and anatomic landmark based statistical models.⁵⁵⁻⁵⁹ These models are not used to consider longitudinal anatomic changes, but only describe the magnitude of random tissue deformation based on a series of patient images acquired over time. The images are considered depictions of random instances of anatomy, but also depict the effects of longitudinal changes. Misattributing longitudinal changes in anatomy to random variation not only misrepresents the magnitude of the latter, but forfeits the utility of recognizing and understanding the former.

In radiation therapy, a predictive model unrestricted by limitations (i) and (ii) can potentially anticipate the manner in which a patient's tumor or normal anatomy changes during treatment. Such a model of morphology has implications for the treatment management of individual patients as well as radiation therapy research studies. Prior to the onset of a particular patient's treatment, a predictive morphology model may be employed to forecast his or her anatomic response to the treatment. This information could be used by clinicians to make better informed decisions regarding the number and timing of treatment plan adjustments or other interventions. A predictive model based on a generative descriptor also allows the forecast morphology to be depicted anatomically. As a result, clinicians can assess the proximity of the target to organs at risk or consider its position within an intended dose distribution. This information may influence geometric and dosimetric decisions during the treatment planning process. The utility of a predictive, generative model continues during the patient's treatment. The model might be used to determine if a particular patient is

responding as expected considering similar patients. If he or she is not, adjustments to the treatment or other interventions can be considered. After completion of a patient's therapy, his or her response may be compiled with those of others to investigate patterns and correlations with additional patient characteristics. This information could advance understanding of patient-specific treatment responses, leading to improved treatment management.

This work focuses on predictive models unrestricted by limitations (i) and (ii). Models based on the reductive descriptors of tumor volume and centroid position, as well as some based on generative descriptions of tumor morphology, are used to describe the dynamic changes of the tumor during the course of treatment. The predictive accuracies of numerous models are evaluated. The models are then compared and the potential benefit of using those of a certain form and complexity is determined. This purpose is divided into four specific aims designed to support a principal hypothesis.

The remainder of this dissertation is organized as follows: Chapter 2 presents the principal hypothesis and specific aims of the work; Chapter 3 describes the methodology used to conduct the work for each of the four specific aims; Chapter 4 provides figures and summaries of the results of the methodology; Chapter 5 features discussion of these results; and Chapter 6 presents the conclusions of this work.

Chapter 2

PRINCIPAL HYPOTHESIS AND SPECIFIC AIMS

Principal Hypothesis:

Longitudinal trends in the volume, position, and morphology of head and neck tumors can be derived from previous patients and applied to reduce the error in forecasting these variables during the treatment of a new patient by 10% compared to the original planning CT.

Specific Aim 1: To generate, evaluate, and compare models that predict the longitudinal changes in *GTV volume and GTV centroid position* throughout the course of radiation therapy, and to discern the effect of *the frequency with which model parameters are adjusted*.

Hypothesis: Models of the longitudinal changes in *GTV volume and GTV centroid position* with *model parameter adjustments* can reduce the error in forecasting these variables during treatment of a new patient by 10% compared to models based exclusively on the planning CT.

Specific Aim 2: To generate, evaluate, and compare models that predict the longitudinal changes in *GTV morphology* throughout the course of radiation therapy, and to discern the effect of *the frequency with which model parameters are adjusted*.

Hypothesis: Models of the longitudinal changes in *GTV morphology* with *model parameter adjustments* can reduce the error in forecasting this variable during treatment of a new patient by 10% compared to models based exclusively on the planning CT.

Specific Aim 3: To generate, evaluate, and compare models that predict the longitudinal changes in *GTV volume and GTV centroid position* throughout the course of radiation therapy, and to discern the effect of *the type of model selected*.

Hypothesis: Non-static models of the longitudinal changes in *GTV volume and GTV centroid position* of previous patients can reduce the error in forecasting these variables during treatment of a new patient by 10% compared to the initial values of the planning CT.

Specific Aim 4: To generate, evaluate, and compare models that predict the longitudinal changes in *GTV morphology* throughout the course of radiation therapy, and to discern the effect of *the type of model selected*.

Hypothesis: Non-static models of the longitudinal changes in *GTV morphology* of previous patients can reduce the error in forecasting this variable during treatment of a new patient by 10% compared to the initial values of the planning CT.

Table I: Illustration of the relationship between the four Specific Aims.

<u>Specific Aim Formula:</u> To generate, evaluate, and compare models that predict the longitudinal changes in (A) throughout the course of radiation therapy, and to discern the effect of (B)		(A)	
		...GTV volume and GTV centroid position...	...GTV morphology...
(B)	...the frequency with which model parameters are adjusted.	Specific Aim 1	Specific Aim 2
	...the type of model selected.	Specific Aim 3	Specific Aim 4

*Chapter 3***METHODOLOGY****3.1. Instances of the gross tumor volume and deformable image registration**

Nineteen patients previously treated for oropharyngeal cancer were randomly selected from an adaptive radiation therapy protocol described by Schwartz et al.³⁶ Patient selection criteria included stage III, IVa, or IVb squamous cell carcinoma of the oropharynx according to the American Joint Committee on Cancer or a status of 0 – 2 according to the Eastern Cooperative Oncology Group. Patient information is presented in Table II. The prescription dose for each patient was 69.96 Gy or 70.00 Gy delivered in 2.12 Gy or 2.00 Gy fractions. Per the clinical protocol, patients had received daily CT-on-rails imaging for image-guided alignment. The CT-on-rails system is depicted in Figure 1. It consists of a 2100 EX linear accelerator (Varian Medical Systems, Palo Alto, CA) and a SmartGantry CT (GE Healthcare, Waukesha, WI) which share a single patient couch. At each treatment fraction, prior to the delivery of radiation, the patient was positioned on the couch and imaged by the CT. The couch was then rotated 180 degrees into position with respect to the linear accelerator. The daily CT image set that had just been acquired was then used to calculate any necessary patient positioning corrections using the C2 vertebral body as the alignment target. Once the patient was properly positioned, the radiation was delivered. Precision of the CT-on-rails system is described by Court et al.⁶⁰

Table II: Patient information

Patient	Age (Years)	Disease Site	Disease Stage*	Concurrent Systemic Therapy
1	62	Base of Tongue	T1 N2A	Cisplatin
2	41	Tonsil	T3 N2B	Cisplatin & Carboplatin
3	53	Tonsil	T2 N2C	Cetuximab
4	56	Base of Tongue	T3 N2C	Cisplatin, Carboplatin, & Paclitaxel
5	57	Base of Tongue	T3 N2A	Cisplatin
6	58	Base of Tongue	T4 N2B	Cisplatin & Cetuximab
7	50	Base of Tongue	T2 N2B	Cetuximab
8	49	Tonsil	T1 N2B	Cisplatin
9	42	Base of Tongue	T2 N2A	Cisplatin
10	54	Base of Tongue	T4 N0	Cisplatin
11	62	Tonsil	T4 N0	Cetuximab
12	56	Base of Tongue	T2 N2B	Cisplatin
13	50	Base of Tongue	T2 N2A	Cisplatin
14	51	Tonsil	T3 N2B	Cisplatin
15	50	Base of Tongue	T4 N2B	Cisplatin
16	38	Tonsil	T1 N2B	Cisplatin
17	68	Tonsil	T2 N2B	Cisplatin
18	69	Tonsil	T3 N1	Cisplatin
19	44	Base of Tongue	T2 N2B	Cetuximab

*No patient had metastatic disease.



Figure 1: CT-on-rails system

On the planning CT images of these 19 patients, 35 gross tumor volumes (GTVs) had been identified and delineated by consensus of several radiation oncologists. These 35 GTVs consisted of 17 primary GTVs and 18 nodal GTVs. Tracking changes of the GTVs that occurred over the course of treatment required contours on each set of daily CT images in addition to those on the planning CT images. Deformable image registration (DIR) was used to facilitate the generation of these additional contours. For each patient, after an initial rigid registration, an in-house DIR algorithm was used to register each set of daily CT images to the patient's planning CT images. The vector field resulting from each registration described the mathematical transformation between the two image sets and was used to propagate the contours from the planning CT images onto the daily CT images. Each deformed contour was reviewed individually and residual inaccuracies were considered acceptable for the analysis. For the sake of consistency, the first 32 treatment fractions were considered even though some patients received as many as 35 treatment fractions. A 20th patient had been considered initially but was omitted due to an insufficient number of available daily CTs.

The DIR algorithm was based on Thirion's demons algorithm⁶¹ and improved by Wang et al.⁶² The demons algorithm operates as if agents or *demons* located at each voxel were applying forces to deform a moving image (m) with velocity (v) based on the gradient of the difference between it and a static image (s). This is expressed mathematically in Equation 1.

$$\vec{v} = \frac{(m-s)\vec{\nabla}s}{(\vec{\nabla}s)^2 + (m-s)^2} \quad (1)$$

Wang et al. incorporated a second term (Equation 2) which swaps the roles of the original moving and static images to improve accuracy and efficiency to describe the force (f) applied to deform the image.

$$\vec{f} = \frac{(m-s)\vec{v}s}{(\vec{v}s)^2 + (m-s)^2} - \frac{(s-m)\vec{v}m}{(\vec{v}m)^2 + (s-m)^2} \quad (2)$$

Wang et al. also assessed the accuracy of the algorithm by applying it to head and neck anatomy deformed by a known transformation. The average error between the results of the algorithm and the known transformation was 0.2 mm (standard deviation: 0.6 mm) while 99% of the errors were within 2 mm.⁶³ Performance of this algorithm has also been compared with a number of other DIR algorithms and performed favorably.⁶⁴

3.2. Descriptor variables

In this work, four variables were used to describe the GTVs. Two of the variables, GTV volume and GTV centroid position, were reductive descriptors, while the other two were generative descriptors that represented GTV morphology. Each descriptor variable was converted to a feature vector that was used for the generation of predictive models. The descriptor variables and their conversion into feature vectors are expounded below.

3.2.1. GTV volume descriptor

The contours of each GTV in each CT image set were converted to a triangular surface mesh. The encompassed volume was then discretized into 1 mm isotropic voxels. The volume of the GTV was calculated as the product of the total number of voxels and the voxel volume (1 mm³). Although volumetric information is directly available from the treatment planning system (Pinnacle3, Philips Medical Systems, Andover, MA), it was determined in the manner described here as part of the more comprehensive image processing workflow. For each GTV in each CT image set, a single-element feature vector was generated from the numeric volume of the GTV.

3.2.2. Bony anatomy landmarks

The three remaining descriptor variables, the GTV centroid position descriptor and the two GTV morphology descriptors, were based on the locations of one or more points in the patients' anatomy. The locations of these points were considered relative to patient-specific, bony-anatomy coordinate systems in order to establish a frame of reference consistent across patients and CT image sets.

The bony-anatomy coordinate systems were created by identifying the location of four bony landmarks on each set of CT images. The landmarks were 1) the basion, at the midpoint on the anterior margin of the foramen magnum 2) the incisive foramen, midline in the anterior hard palate, 3) the left cochlea, and 4) the right cochlea. Figure 2 depicts the anatomic location of these four landmarks. The landmarks were identified on a radiation therapy treatment planning system using a "bone" window and level setting. A single viewer positioned the landmarks, preventing inter-observer variability.

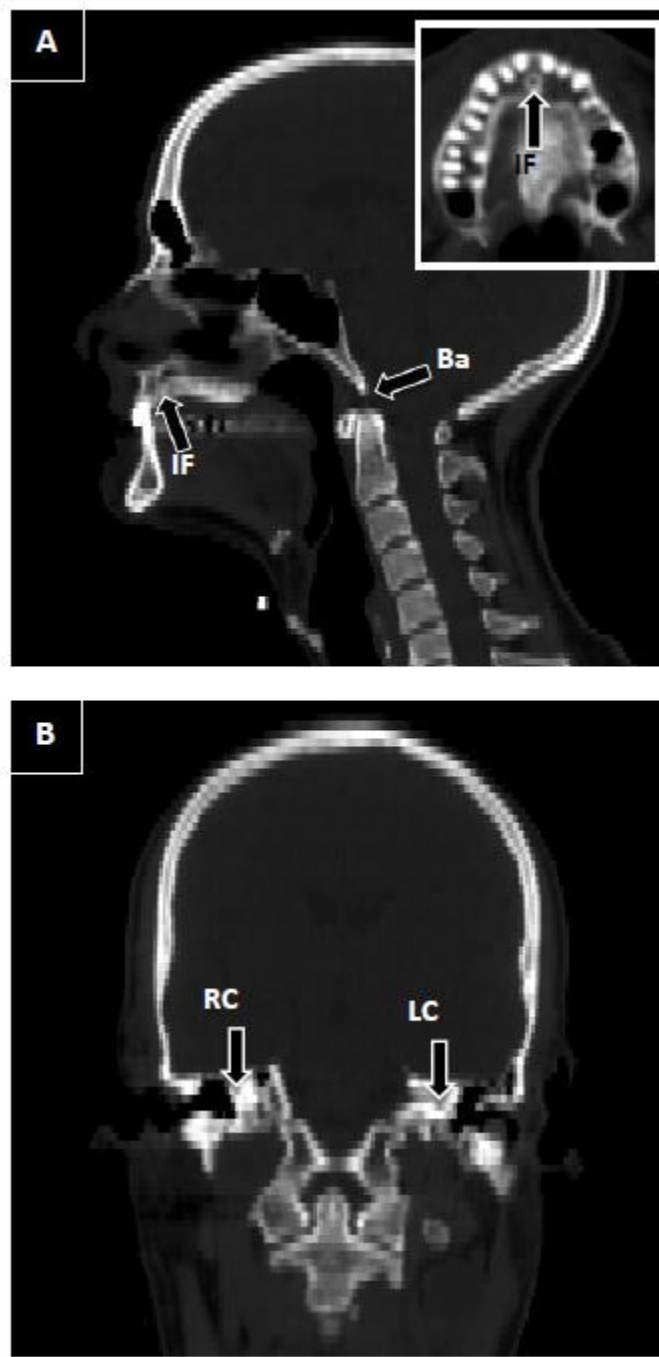


Figure 2: Locations of bony anatomy landmarks. (A) Sagittal plane depicting the basion (Ba) and the incisive foramen (IF) landmarks. Insert: Close-up of incisive foramen landmark in axial plane. (B) Coronal plane depicting the right cochlea (RC) and the left cochlea (LC) landmarks.

The lateral (LR) dimension of the bony-anatomy coordinate system was defined parallel to the line between the left and right cochlea landmarks. The anterior-posterior (AP) dimension was initially defined parallel to the line between the basion and incisive foramen landmarks. The cross product between the LR and AP dimensions defined the superior-inferior (SI) dimension. The SI dimension was crossed with the LR dimension to re-define the AP dimension and ensure orthogonality. The bony-anatomy coordinate system was shifted to position its origin at the midpoint between the left and right cochlea landmarks.

With no ground truth regarding the correct position of the bony landmarks, it is impossible to measure the error in their placement. The precision, or reproducibility, of their placement, however, was calculated from inter-landmark distances. The propagation of uncertainties in individual landmark placements to uncertainties in inter-landmark distances are presented in Equations 3 – 6 and in an equivalent matrix format in Equation 7. The uncertainty in the position of a point i or the distance between points i and j are denoted by σ_i and σ_{ij} , respectively. A and B denote the basion and incisive foramen landmarks, respectively, and C and D denote the cochlea landmarks interchangeably. The uncertainties in placing each bony anatomy landmark was calculated by solving Equation 7 using ordinary least squares optimization.

$$\sigma_{AC}^2 = \sigma_A^2 + \sigma_C^2 = \sigma_A^2 + \sigma_D^2 = \sigma_{AD}^2 \quad (3)$$

$$\sigma_{BC}^2 = \sigma_B^2 + \sigma_C^2 = \sigma_B^2 + \sigma_D^2 = \sigma_{BD}^2 \quad (4)$$

$$\sigma_C^2 = \sigma_D^2 \quad (5)$$

$$\sigma_{CD}^2 = \sigma_C^2 + \sigma_D^2 = 2\sigma_C^2 \quad (6)$$

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} \sigma_A^2 \\ \sigma_B^2 \\ \sigma_C^2 \end{bmatrix} = \begin{bmatrix} \sigma_{AC}^2 \\ \sigma_{BC}^2 \\ \sigma_{CD}^2 \end{bmatrix} \quad (7)$$

Assumed to be isotropic, the uncertainties were subsequently used to calculate the uncertainty in the position (σ_{origin}) and orientation (σ_{pitch} , σ_{roll} , and σ_{yaw}) of the bony-anatomy coordinate system (Equations 8 – 10). d_{ij} denotes the average distance between points i and j .

$$\sigma_{origin} = \frac{\sigma_C}{\sqrt{2}} \quad (8)$$

$$\sigma_{pitch} = \frac{\sqrt{\sigma_A^2 + \sigma_B^2}}{d_{AB}} \quad (9)$$

$$\sigma_{roll} = \sigma_{yaw} = \frac{\sqrt{\sigma_C^2 + \sigma_D^2}}{d_{CD}} \quad (10)$$

3.2.3. GTV centroid position descriptor

The position of the geometric centroid of the GTV depicted on each CT image set was determined by the treatment planning system. A feature vector was created as a row vector containing the lateral, anterior-posterior, and superior-inferior coordinates of the GTV centroid relative to the bony-anatomy coordinate system.

3.2.4. Radial extent GTV morphology descriptor

Two generative descriptors were used to represent GTV morphology. Both descriptors were based on landmarks placed on the surface of each GTV. A computer algorithm positioned these surface landmarks where lines radiating from the GTV centroid intersected the GTV surface mesh. The lines radiating from the GTV centroid were equally spaced at 10 degree polar and azimuthal angles with respect to the bony-anatomy coordinate system. The result was 614 surface landmarks on each GTV.

The first of the two generative descriptors was the radial extent GTV morphology descriptor and was based directly on the set of 614 GTV surface landmarks. The distances between the GTV centroid and the GTV surface landmarks were listed as a row vector. This vector was concatenated with the three coordinates of the GTV centroid to yield a 617-element feature vector. This closely resembled the morphometric analysis technique of Lele and Richtsmeier in which a configuration of landmarks is described by the complete set of inter-landmark distances.⁶⁵ In the work presented here, only a select subset of inter-landmark distances was considered.

3.2.5. Spherical harmonic GTV morphology descriptor

The second generative descriptor was the spherical harmonic GTV morphology descriptor. It was based on a spherical harmonic decomposition of the distances between the GTV centroid and the GTV surface landmarks. Spherical harmonics are the angular part of solutions to Laplace's equation represented in spherical coordinates (Equation 11). The real form of spherical harmonics ($Y_l^m(\theta, \varphi)$) is provided in Equation 12 and depicted in Figure 3. The degree of the spherical harmonic (l) is a whole number, the order of the spherical harmonic (m) is an integer between $-l$ and l , inclusive, and $P_l^m(\cdot)$ represents the associated Legendre Polynomials. Polar and azimuthal coordinates are denoted θ and φ , respectively.

$$\Delta f = \frac{1}{\rho^2} \frac{\delta}{\delta \rho} \left(\rho^2 \frac{\delta f}{\delta \rho} \right) + \frac{1}{\rho^2 \sin \theta} \frac{\delta}{\delta \theta} \left(\sin \theta \frac{\delta f}{\delta \theta} \right) + \frac{1}{\rho^2 \sin^2 \theta} \frac{\delta^2 f}{\delta \varphi^2} = 0 \quad (11)$$

$$Y_l^m(\theta, \varphi) = \begin{cases} \sqrt{\frac{2l+1}{2\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos \theta) \cos(m\varphi) & \text{if } m > 0 \\ \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos \theta) & \text{if } m = 0 \\ \sqrt{\frac{2l+1}{2\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos \theta) \sin(m\varphi) & \text{if } m < 0 \end{cases} \quad (12)$$

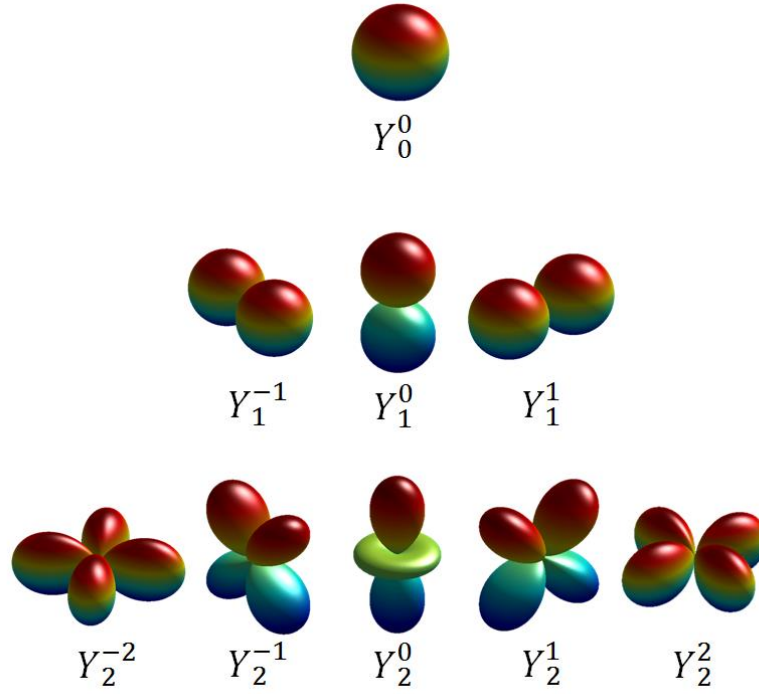


Figure 3: Spherical harmonics. Y_l^m is depicted for degree $0 \leq l \leq 2$ and order $-2 \leq m \leq 2$.

Frequently, a set of spherical harmonic terms are used as an orthogonal basis to quantify the notion of shape. Applications of this type have been used to describe diverse phenomena such as atomic orbitals,^{66, 67} molecular binding sites,⁶⁸ computer graphics lighting,^{69, 70} planetary gravitational and magnetic fields,⁷¹⁻⁷⁴ and the cosmic microwave background.⁷⁵ A representation of spherical harmonics used as a basis is depicted in Equation 13. Here α_l^m denotes the set of unknown coefficients weighing each spherical harmonic term. Projecting a particular shape onto the spherical harmonic basis is equivalent to solving for α_l^m .

$$f(\theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \alpha_l^m Y_l^m(\theta, \varphi) \quad (13)$$

The configuration of the 614 surface landmarks of each GTV was projected onto the spherical harmonic basis. This was accomplished by recasting Equation 13 as a matrix multiplication and setting it equal to the distances between the GTV centroids and the GTV

surface landmarks ($r(\theta, \varphi)$). Ordinary least squares optimization was then used to solve for α_l^m (Equation 14).

$$[Y_l^m(\theta, \varphi)][\alpha_l^m] = [r(\theta, \varphi)] \quad (14)$$

The spherical harmonics of every order up through degree 9, a total of 100 terms, were included in the basis. The resulting 100 spherical harmonic coefficients were listed in a row vector. This vector was concatenated with the three coordinates of the GTV centroid to yield a 103-element feature vector.

3.2.6. Regularizing feature vector elements

For each GTV, the feature vectors derived from the four descriptor variables changed over the course of treatment. The set of feature vectors pertaining to a particular GTV throughout treatment represented the changing state of the GTV. Each element of these vectors was viewed as sample from a time series. To reinforce consideration of the feature vector elements as continuous functions, each time series was smoothed using a cubic smoothing spline. The parameter of the smoothing spline was chosen qualitatively to fit the data while enforcing a smooth function. The smooth function was sampled to yield new feature vectors for model generation.

3.3. Forecast models

Six models were used to forecast the changes in the descriptor variable feature vectors at each treatment fraction over the course of therapy. They are referred to as the static, linear, mean, median, vector PCA, and 2D PCA models, and each was characterized by a different set of model parameters as described below. The model parameters were determined from training GTVs and applied to test GTVs as described in Section 3.6.

3.3.1. Static model

The first of the models, the static model, assumed no change in the feature vector elements from their initial values depicted in the planning CT images. A forecast according to this model consisted of the un-normalized initial value of the test GTV feature vector perpetuated throughout the duration of treatment. This is represented in Equation 15 where $V_{0,test}$ and $V_{t,test}$ denote the feature vectors corresponding to the planning CT and the daily-acquired CT from treatment fraction t , respectively. Conceptually, this model is similar to the conventional, non-adaptive radiation therapy management strategy where a single treatment plan is optimized based on the planning CT images and administered without adjustment for the entirety of treatment.

$$V_{t,test} = V_{0,test} \quad (15)$$

3.3.2. Linear model

A forecast according to a linear model was the linear interpolation between the test GTV feature vector's initial value and the median relative value of the training GTV feature vectors at the final treatment fraction (treatment fraction 32). These interpolated values were

applied to the initial, un-normalized value of the test GTV to forecast feature vectors for the entire treatment. Equations 16 and 17 describe generation of the linear model. Here, β is the constant rate of change determined by the feature vectors of the training GTVs ($V_{t,train}$).

$$V_{t,test} = V_{0,test} \times (1 + \beta \times t) \quad (16)$$

$$\beta = median\left(\frac{V_{32,train} - V_{0,train}}{V_{0,train}}\right) \times \left(\frac{1}{32}\right) \quad (17)$$

3.3.3. Mean model

A forecast according to the mean model was determined by calculating the mean relative value of the training GTV feature vectors at each treatment fraction (Equation 18). These values were applied to the initial, un-normalized value of the test GTV to forecast feature vectors for the entire treatment.

$$V_{t,test} = V_{0,test} \times mean\left(\frac{V_{t,train}}{V_{0,train}}\right) \quad (18)$$

3.3.4. Median model

Similar to that of the mean model, a forecast according to the median model was created by calculating the median relative value of the training GTV feature vectors at each treatment fraction. These values were applied to the initial, un-normalized value of the test GTV to forecast feature vectors for the entire treatment (Equation 19).

$$V_{t,test} = V_{0,test} \times median\left(\frac{V_{t,train}}{V_{0,train}}\right) \quad (19)$$

Because this model was based on the median rather than the mean, it was less susceptible to the influence of extreme values of the feature vector elements. These outlying

values may represent measurement or calculation errors, or instances of feature vector elements considerably dissimilar to the others.

3.3.5. Vector PCA and 2D PCA models

Two models were based on principal component analysis (PCA) of the feature vectors and are referred to as the vector PCA model and the 2-dimensional (2D) PCA model. PCA is a quantitative analysis technique that identifies the primary modes of variation among the elements of some data vector, in this case, the GTV feature vectors. To identify the primary modes of variation, first, a covariance matrix of the feature vector elements was generated. This was achieved differently for the two PCA models.

The vector PCA model generated the covariance matrix in the manner of conventional PCA. For each GTV, there were 33 instances of the feature vector for each of the four descriptor variables. The instances were of length n (where $n = 1, 3, 617$, or 103 , depending on the variable) and were derived from the planning CT images and the daily CT images acquired at each of the 32 treatment fractions. A $33 \times n$ matrix was created from the stack of feature vectors. This matrix was considered a single observation of the manner in which the GTV feature vectors changed during treatment. Because generation of the covariance matrix operates on a stack of observations as row vectors, each $33 \times n$ matrix needed to be vectorized into a $1 \times 33n$ vector as depicted in Equation 20. The covariance matrix was then calculated from the stack of these row observation vectors.

$$\text{vec} \left(\begin{bmatrix} V_{0,1} & V_{0,2} & \dots & V_{0,n} \\ V_{1,1} & V_{1,2} & \dots & V_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ V_{32,1} & V_{32,2} & \dots & V_{32,n} \end{bmatrix} \right) = [V_{0,1}, V_{0,2}, \dots, V_{0,n}, V_{1,1}, V_{1,2}, \dots, V_{1,n}, V_{32,1}, V_{32,2}, \dots, V_{32,n}] \quad (20)$$

For the 2D PCA model, the covariance matrix was generated without vectorizing the matrices. Instead, the covariance matrix (\mathbf{G}) was calculated from the 2-dimensional observations according to Equation 21 where \mathbf{A}_j is the j^{th} observation matrix.

$$\mathbf{G} = \frac{1}{M} \sum_{j=1}^M (\mathbf{A}_j - \bar{\mathbf{A}})^T (\mathbf{A}_j - \bar{\mathbf{A}}) \quad (21)$$

Two advantages of avoiding the vectorization step were that 1) the covariance matrix was considerably smaller, easing subsequent computation, and 2) structure of data inherently organized as a matrix with correspondence among elements in each column was maintained. Described by Yang et al., this technique is appropriate for inherently 2-dimensional data structures such as images of faces.⁷⁶

After the covariance matrix was generated, it was decomposed into its eigenvectors and associated eigenvalues. The resulting eigenvectors were the principal components (PCs) of variation among the elements of the observations, and the associated eigenvalues represented the variance of the observations along the PCs. The PCs were ranked in importance by ordering them by the magnitude of their variances. The eigenvector with the largest eigenvalue was denoted as the 1st PC, the eigenvector with the second largest eigenvalue was the 2nd PC, and so on. PCA is commonly used to reduce the complexity and dimensionality of data without sacrificing an excessive amount of information. This is achieved by retaining only the most influential PCs and ignoring the rest. Beginning with the 1st PC, additional PCs were incorporated in order of decreasing variance until the sum of included variances exceeded a predetermined threshold proportion of the total variance (90%).

For both the vector PCA and 2D PCA models, the original observations were projected onto the selected subset of PCs after the eigen-decomposition. The median values of the observations along each of the selected PCs were back-transformed to represent an average

observation. The feature vector values corresponding to this average observation were applied to the initial, un-normalized value of the test GTV to forecast feature vectors for the entire treatment. The median PC values were used because they were robust to outliers and because the forecast corresponding to the mean PC values was equivalent to that of the mean model above. The median PC values, however, did not necessarily generate a forecast equivalent to the median model above.

3.4. Representing changes in feature vector elements

The GTV centroid position feature vectors and both morphology feature vectors all included three elements that designated the position of the GTV centroid. Due to the large variation in the initial centroid position, the absolute change of these elements, rather than their relative change, was considered. Changes of the remaining elements were described relative to their initial value. An additional constraint capped the maximum daily change at 25% for all elements of the spherical harmonic feature vector in order to regularize the forecast. This value was determined based on preliminary modeling experience.

3.5. Model adjustment schemes

To explore the effects of incorporating patient-specific information acquired during treatment, each of the models above was evaluated with and without a number of adjustment points. Adjustment points were opportunities for the model parameters, and thus the forecast feature vectors, to be updated to match more recently observed patient-specific values. This mimicked an adaptive radiation therapy treatment where the patient was re-imaged and a new plan was optimized at each adjustment point. Model adjustment schemes incorporating 0, 1,

2, 3, or 5 adjustment points were compared. This included multiple schemes with 2 or 3 adjustments positioned at various treatment fractions. A total of 11 different adjustment schemes were considered and are illustrated in Figure 4.

Time Points:	0	1	6	11	16	21	26	31
Number of Adjustment Points:	0	●	○ ○ ○ ○ ○	○ ○ ○ ○ ○	○ ○ ○ ○ ○	○ ○ ○ ○ ○	○ ○ ○ ○ ○	○ ○
	1	●	○ ○ ○ ○ ○	● ○ ○ ○ ○	○ ○ ○ ○ ○	○ ○ ○ ○ ○	○ ○ ○ ○ ○	○ ○
		●	○ ○ ○ ○ ○	○ ○ ○ ○ ○	● ○ ○ ○ ○	○ ○ ○ ○ ○	○ ○ ○ ○ ○	○ ○
		●	○ ○ ○ ○ ○	○ ○ ○ ○ ○	○ ○ ○ ○ ○	● ○ ○ ○ ○	○ ○ ○ ○ ○	○ ○
		●	○ ○ ○ ○ ○	○ ○ ○ ○ ○	○ ○ ○ ○ ○	○ ○ ○ ○ ○	● ○ ○ ○ ○	○ ○
	2	●	○ ○ ○ ○ ○	● ○ ○ ○ ○	○ ○ ○ ○ ○	● ○ ○ ○ ○	○ ○ ○ ○ ○	○ ○
		●	○ ○ ○ ○ ○	○ ○ ○ ○ ○	○ ○ ○ ○ ○	● ○ ○ ○ ○	○ ○ ○ ○ ○	○ ○
		●	○ ○ ○ ○ ○	○ ○ ○ ○ ○	○ ○ ○ ○ ○	○ ○ ○ ○ ○	● ○ ○ ○ ○	○ ○
	3	●	○ ○ ○ ○ ○	● ○ ○ ○ ○	○ ○ ○ ○ ○	● ○ ○ ○ ○	○ ○ ○ ○ ○	○ ○
	5	●	○ ○ ○ ○ ○	● ○ ○ ○ ○	● ○ ○ ○ ○	● ○ ○ ○ ○	● ○ ○ ○ ○	○ ○

Figure 4: Diagram of model adjustment schemes. Time point 0 represents the time of the pre-treatment CT. Subsequent time points mark treatment fractions. Model parameters were determined from the pre-treatment CT images and images taken at the adjustment points (closed circles) and applied to subsequent treatment fractions (open circles) until the next adjustment point or the end of treatment. For readability, adjustment schemes are grouped vertically by number of adjustment points.

At each adjustment point, the forecast according to each model was adjusted to match the actual values of the test GTV feature vector at that point. Equations 22 – 26 are similar to Equations 15 – 19 but represent the feature vectors predicted between adjustment points at treatment fractions p and q . These equations also apply to the interval prior to the first adjustment point and subsequent to the last one when $p = 0$ and $q = 32$, respectively.

$$V_{p \leq t < q, test} = V_{p, test} \quad (22)$$

$$V_{p \leq t < q, test} = V_{p, test} \times (1 + \beta_p \times t) \quad (23)$$

$$\beta_p = \text{median} \left(\frac{V_{q, train} - V_{p, train}}{V_{p, train}} \right) \times \left(\frac{1}{q - p} \right) \quad (24)$$

$$V_{p \leq t < q, test} = V_{p, test} \times \text{mean} \left(\frac{V_{t, train}}{V_{p, train}} \right) \quad (25)$$

$$V_{p \leq t < q, test} = V_{p, test} \times \text{median} \left(\frac{V_{t, train}}{V_{p, train}} \right) \quad (26)$$

To elaborate, in order to predict the feature vectors for treatment fractions between to adjustment points according to the static model (Equation 22), the test GTV feature vector at the most recent adjustment point (p) was sustained until the next adjustment point (q). For the linear model, the adjusted forecast was the linear interpolation between the value of the test GTV feature vector at p and the median value of the training GTV feature vectors at q (Equations 23 and 24). The mean and median models were shifted at p so the average change in the feature vector was applied to the observed test GTV value at p rather than to the average of the training GTV feature vectors (Equations 25 and 26). For both the vector PCA and 2D PCA models, the PCs were determined on each inter-adjustment point period. The matrices were thus $f \times n$ where f was the number of treatment fractions between adjustment points. Figure 5 is an example of how the static and linear models were adjusted at an adjustment point mid-treatment.

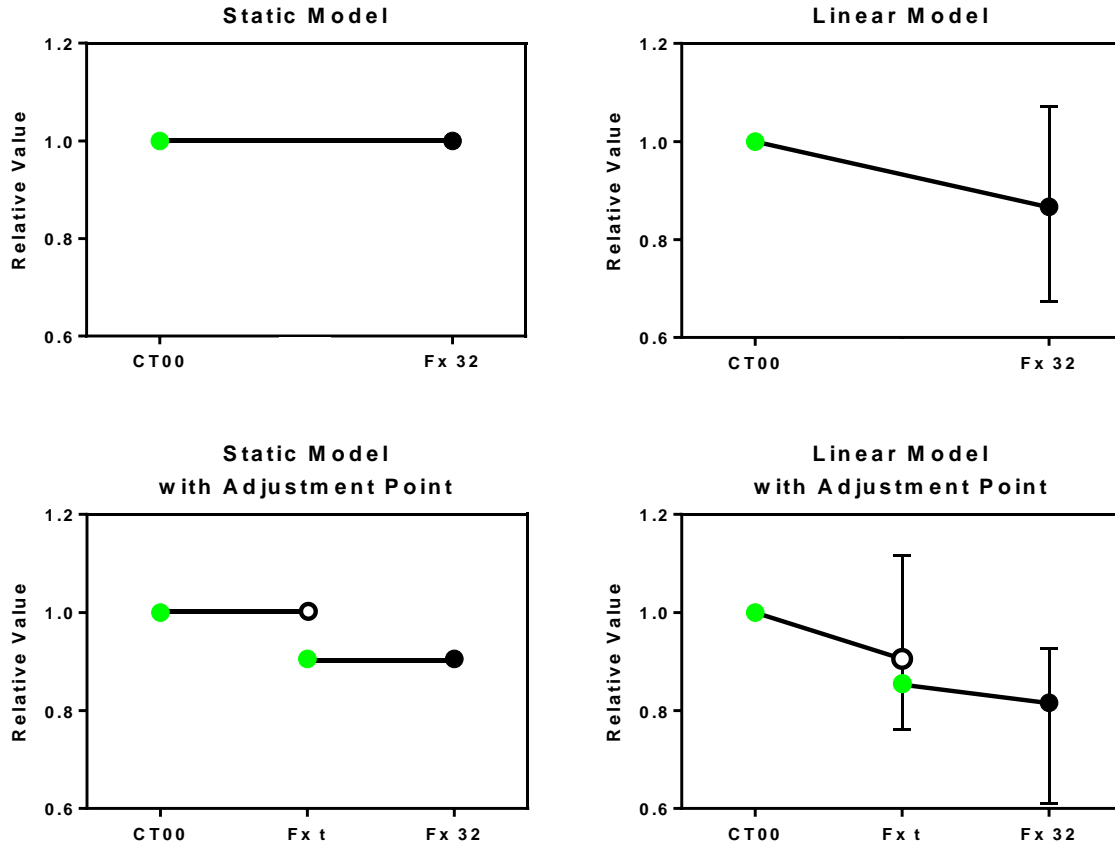


Figure 5: Illustration of static and linear models with and without adjustment points. Green points denote observations and black lines denote forecast values. Error bars represent observed values from the training dataset.

3.6. Evaluation of model forecast accuracy with leave-one-out cross-validation

Evaluation of the forecast accuracy of each model and each model adjustment scheme was conducted using leave-one-out cross-validation. In leave-one-out cross-validation, 1 of the 35 GTVs was identified as the test GTV, and the remaining 34 were considered the training GTVs. The model parameters were established and adjusted as described above based on the normalized feature vectors of the 34 training GTVs. The models were then applied to the un-normalized, initial values of the test GTV to forecast feature vectors at

subsequent treatment fractions. The un-normalized values of the test GTV feature vectors at adjustment points were considered to adjust the model parameters as described above. The accuracy of each model was assessed by calculating the forecast error as defined for each descriptor variable below. This process was repeated 34 times so that each GTV was considered the test GTV once, with an associated forecast error for each of the models.

For the GTV volume descriptor, forecast error was defined as the root mean squared error (RMSE) between the true and forecast test GTV volume at each fraction. For the GTV centroid position descriptor, forecast error was the RMSE between the true and forecast position of the test GTV centroid relative to the bony-anatomy coordinate system at each fraction. Lastly, for the two morphology descriptors, forecast error was the RMSE between the position of the 614 surface landmarks according to the true and forecast test GTV feature vectors at each fraction.

3.7. Statistical comparison of model forecast accuracy

For each of the six model types, the adjustment scheme that resulted in the smallest median forecast error for a particular number of adjustment points was determined. This was trivial for schemes with 0, 3, or 5 adjustment points where there was only one scheme per model type. However, there were five schemes per model type that had 1 adjustment point and three schemes per model type that had 2 adjustment points. The resulting models were considered the most accurate models for a certain level of effort and intervention as represented by the number of adjustment points. Statistical comparisons were made between these models to address Specific Aims 1 – 4. All comparisons were made using two-tailed, Wilcoxon signed-rank tests considered to be significant at $p < 0.05$.

The effect on forecast accuracy of adding adjustment points to models of the reductive descriptor variables (GTV volume and GTV centroid position) was evaluated for Specific Aim 1. This was accomplished by comparing each model of a certain type and number of adjustment points to the model of the same type and no adjustment points. In addition, the marginal effect of adding adjustment points was determined by comparing each model of a certain type and number of adjustment points to the model of the same type and one fewer adjustment points. Both types of comparisons were performed using models of the radial extent and spherical harmonic GTV morphology descriptor variables for Specific Aim 2.

For Specific Aim 3, the effect of model selection on forecast accuracy for the reductive descriptor variables was evaluated. This model type effect was determined by comparing the forecast accuracy of each non-static model to that of the static model with an equal number of adjustment points. The same comparison was made using models of the two generative morphology descriptors for Specific Aim 4.

Chapter 4

RESULTS

4.1. Generating feature vectors

With the exception of the GTV volume descriptor, determination of descriptor variable values depended on a coordinate system based on bony anatomy landmarks. The uncertainty in identifying each bony anatomy landmark was calculated using equations that described how these uncertainties propagated to yield measurable variations in inter-landmark distances (Equations 8 – 10). The resulting uncertainty in identifying the basion, incisive foramen, left cochlea, and right cochlea was 1.1 mm, 1.1 mm, 0.3 mm, and 0.3 mm, respectively. These uncertainties, in turn, lead to a 0.2 mm precision in determining the origin of the bony anatomy coordinate system. The precision of the orientation of the coordinate system was 1.1 degrees (pitch), 0.3 degrees (roll), and 0.3 degrees (yaw).

In order to generate feature vectors for the spherical harmonic GTV morphology descriptor, the number of terms to be included in the spherical harmonic basis needed to be determined. Figure 6 presents the error between the observed position of GTV surface landmarks and their position when modeled using a spherical harmonic basis of increasing degree. This error from truncating the series of spherical harmonic terms decreased with an increase in the degree (and the number of terms). However, the improvement diminished with the addition of subsequent terms. A spherical harmonic basis of degree 9, corresponding to 100 terms, reduced the median error to 0.4 mm and the maximum error to 0.7 mm.

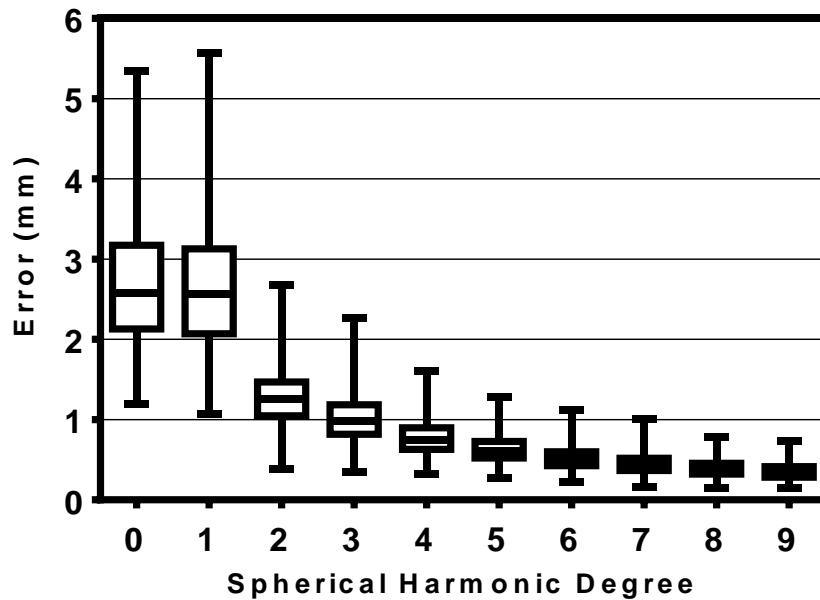


Figure 6: Morphology approximation error according to the maximum degree of the spherical harmonic basis. Box range is 25th to 75th percentile; whisker range is minimum to maximum.

The observed data for each descriptor variable underwent smoothing to discern a functional description of its change over time. The function resulting from applying cubic smoothing splines to the observed data represented a trade-off between generating a smooth function and fitting the data precisely. The residual errors in fitting the data are presented in Figures 7 – 9. These figures depict the root mean squared error (RMSE) between the smoothing spline function and the data of the descriptor variable for each GTV.

For the GTV volume descriptor (Figure 7), the median RMSE was 3.0 mm. For the GTV centroid position descriptor, the median RMSE in the lateral, anterior-posterior, and superior-inferior coordinates were 0.7 mm, 1.3 mm, and 0.7 mm, respectively (Figure 8). The spatial distribution of the smoothing spline error for the radial extent GTV morphology descriptor is illustrated in Figure 9. The size and position of the points in this figure reflect the GTV

surface landmarks depicted with perspective on the surface of a sphere. Their color represents the sub-millimeter residual error from the smoothing process. A depiction of this error for the spherical harmonic morphology descriptor is not presented as the coefficient values represent geometrical constructs rather than intuitive values with units consistent to the original measured data.

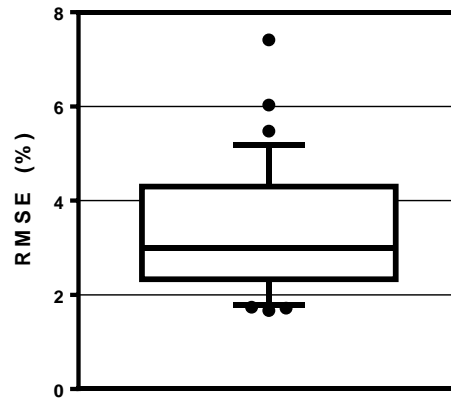


Figure 7: Smoothing spline error – GTV volume descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

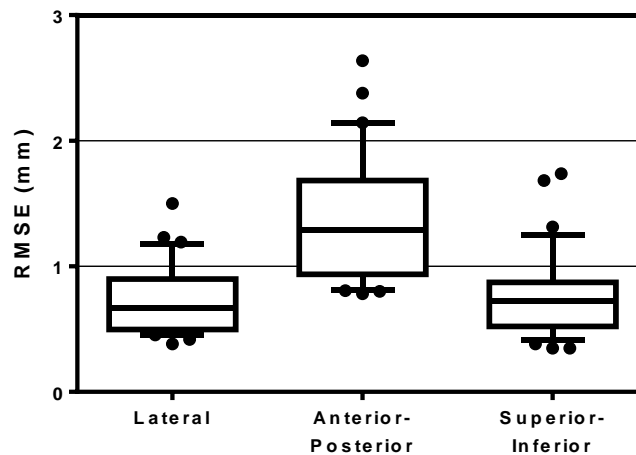


Figure 8: Smoothing spline error – GTV centroid position descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

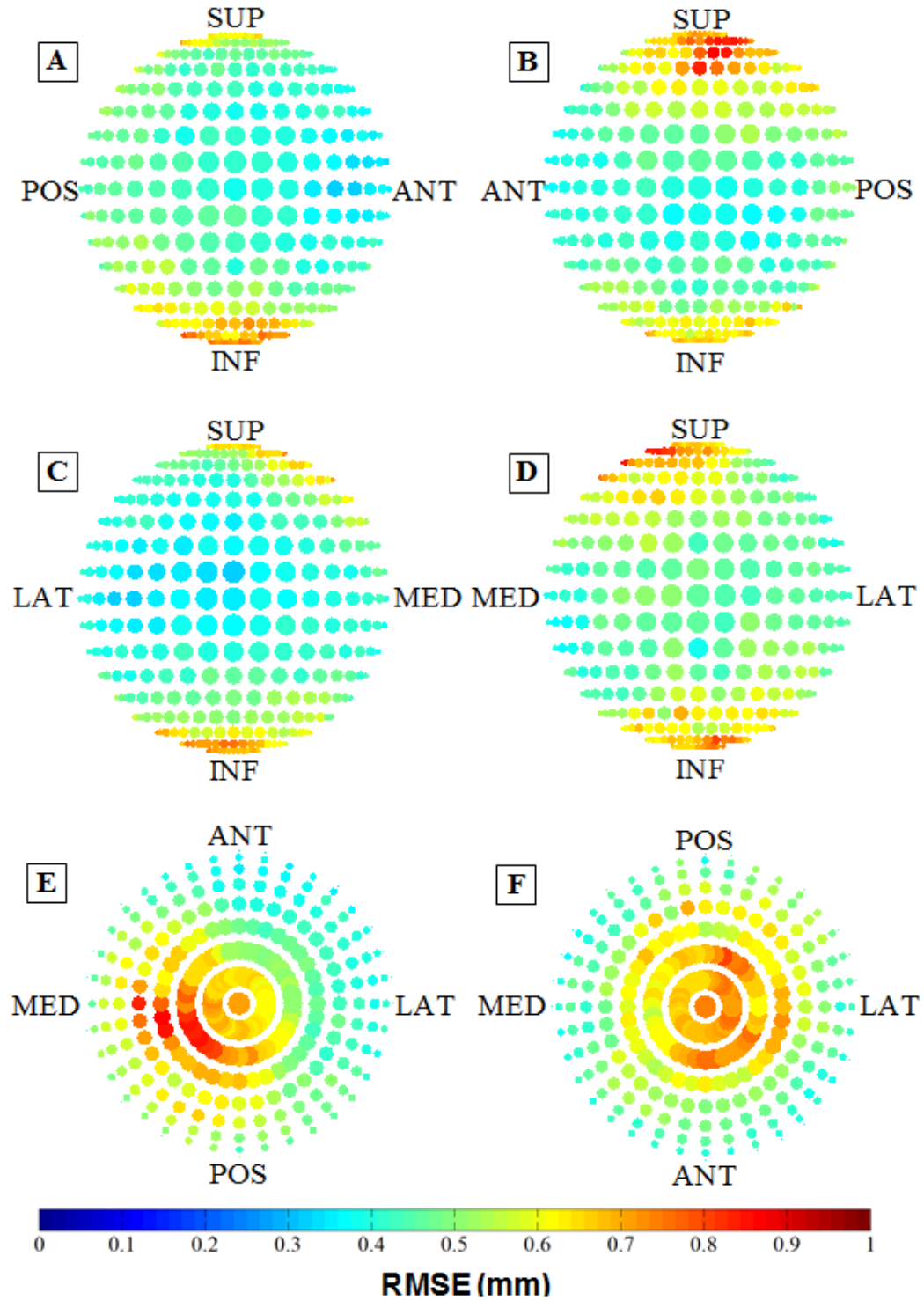


Figure 9: Smoothing spline error – Radial extent GTV morphology descriptor. Points represent surface landmarks on the surface of a sphere viewed from six perspectives: (A) medial (MED), (B) lateral (LAT), (C) posterior (POS), (D) anterior (ANT), (E) inferior (INF), and (F) superior (SUP).

4.2. Results for Specific Aims

The forecast errors of the non-static models of each descriptor variable with different numbers of adjustment points are depicted alongside those of the static models in Figures 10 – 29. Analyzing different combinations of the models elucidated the influence of the number of adjustment points and the type of model on forecast accuracy.

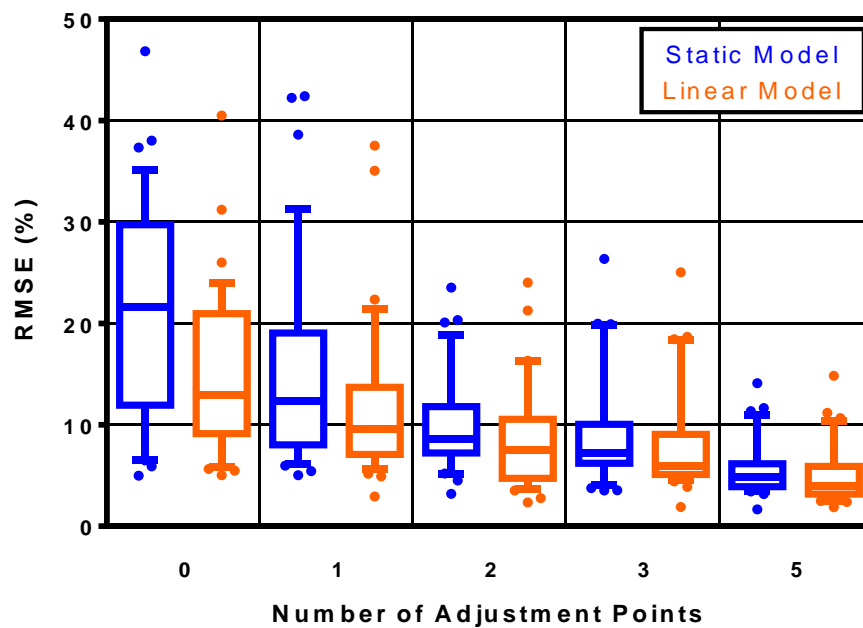


Figure 10: Forecast error of static and linear models – GTV volume descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

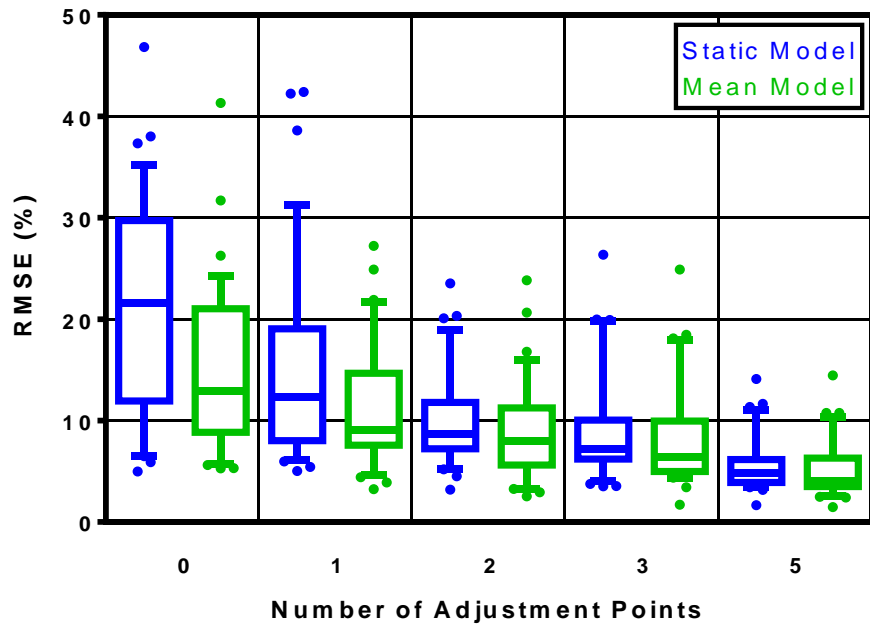


Figure 11: Forecast error of static and mean models – GTV volume descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

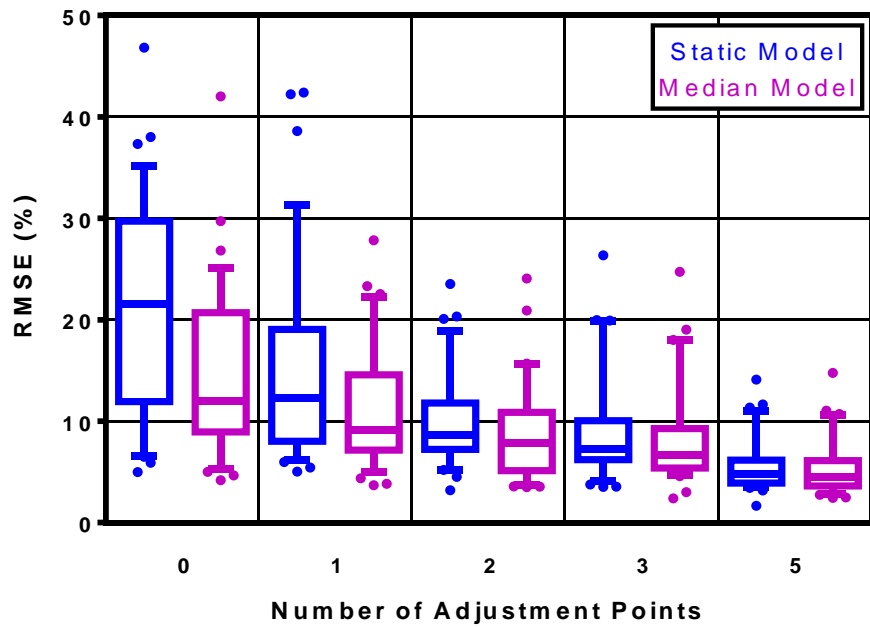


Figure 12: Forecast error of static and median models – GTV volume descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

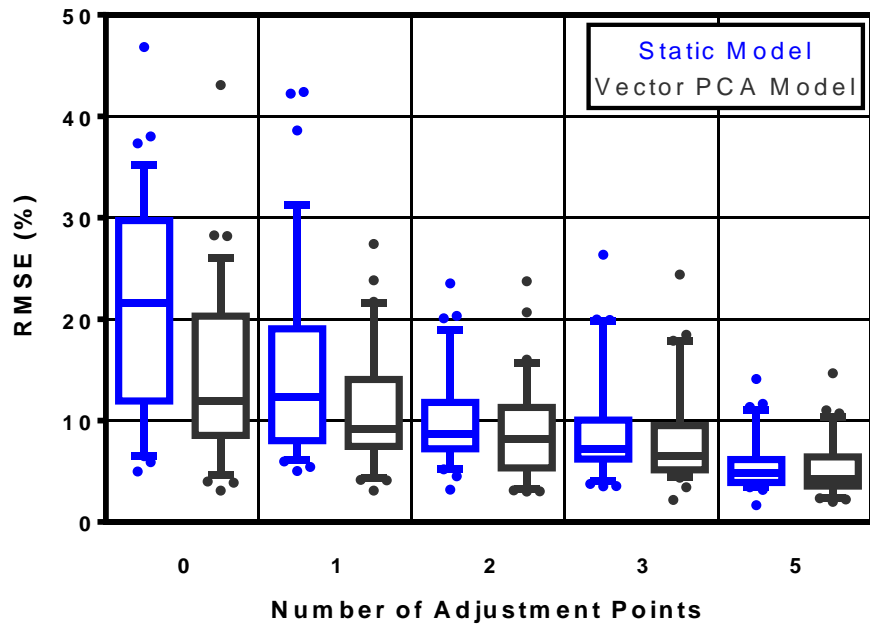


Figure 13: Forecast error of static and vector PCA models – GTV volume descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

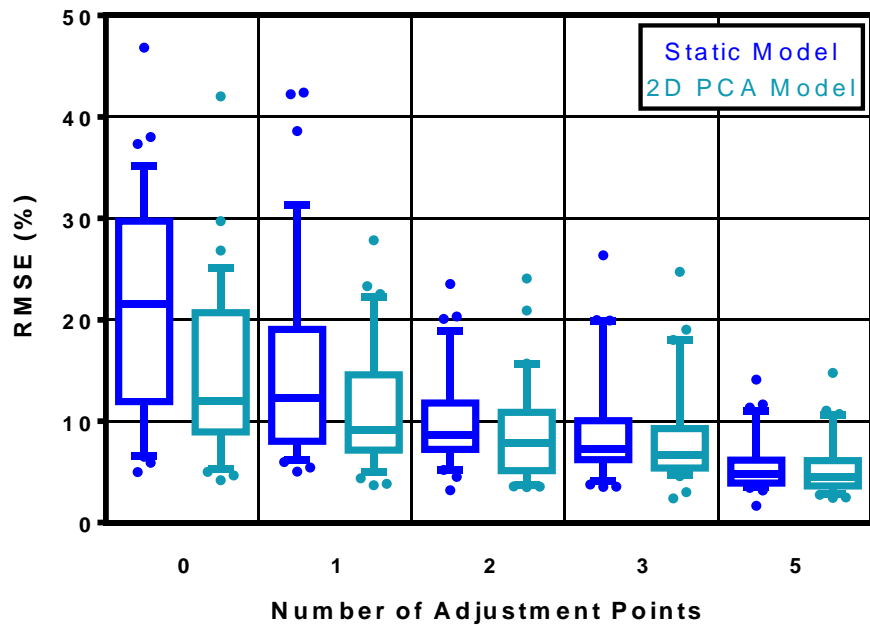


Figure 14: Forecast error of static and 2D PCA models – GTV volume descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

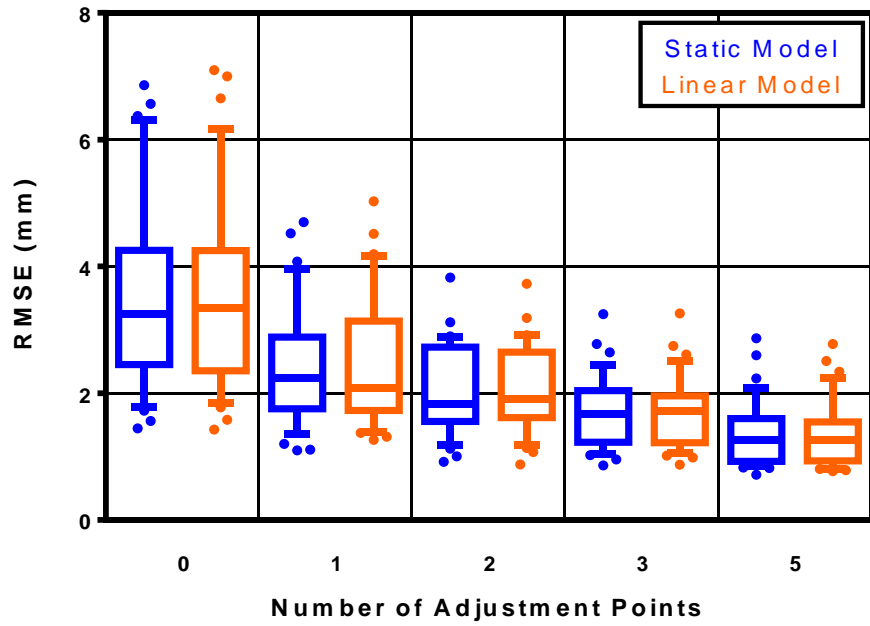


Figure 15: Forecast error of static and linear models – GTV centroid position descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

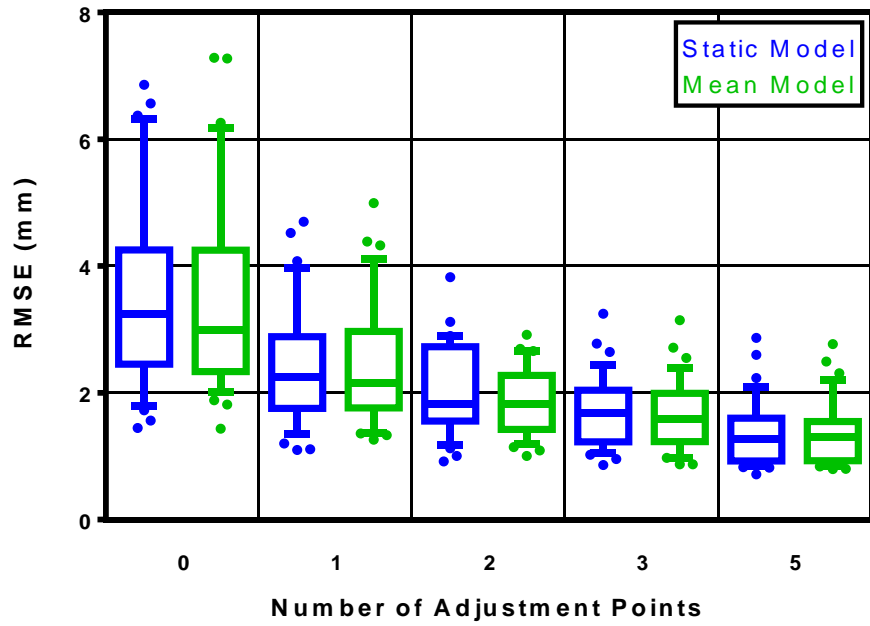


Figure 16: Forecast error of static and mean models – GTV centroid position descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

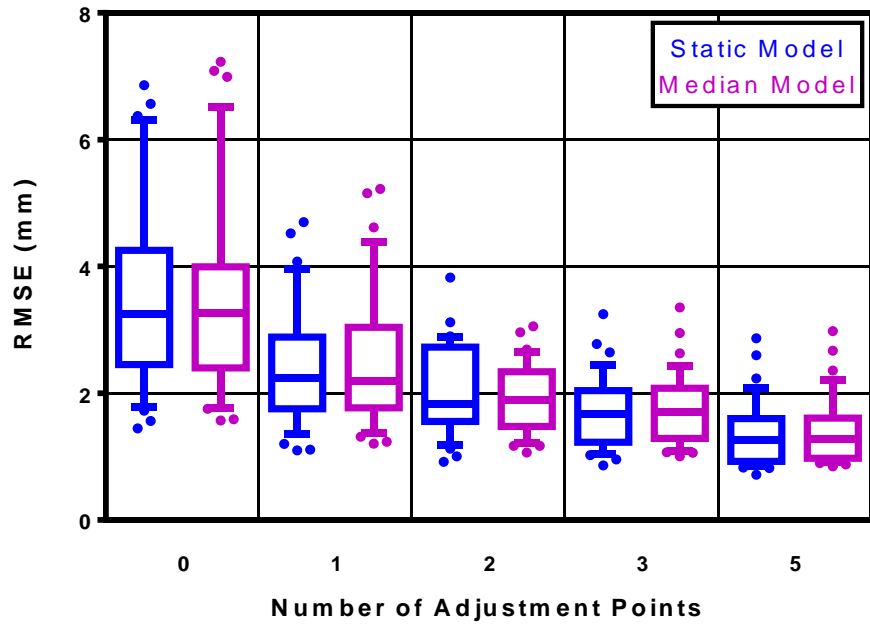


Figure 17: Forecast error of static and median models – GTV centroid position descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

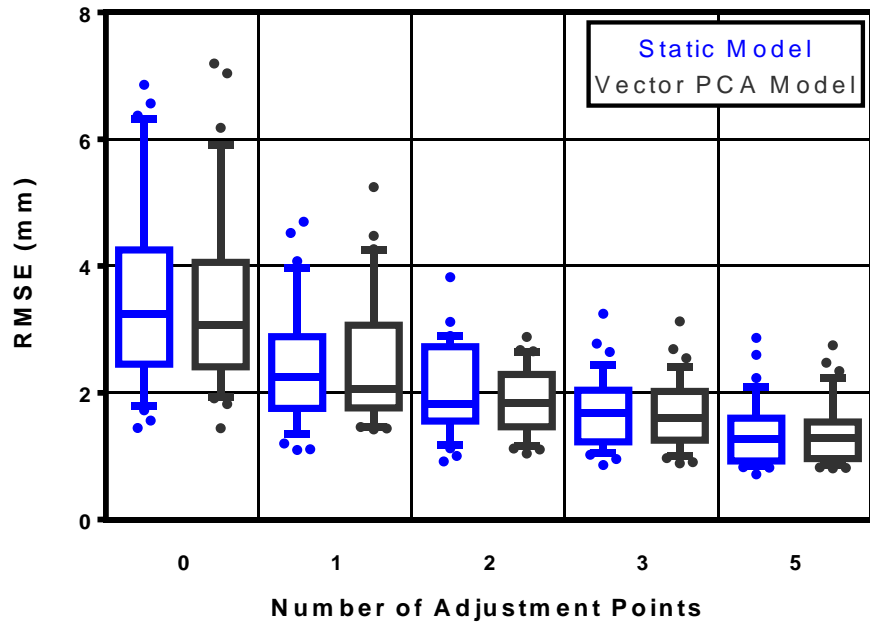


Figure 18: Forecast error of static and vector PCA models – GTV centroid position descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

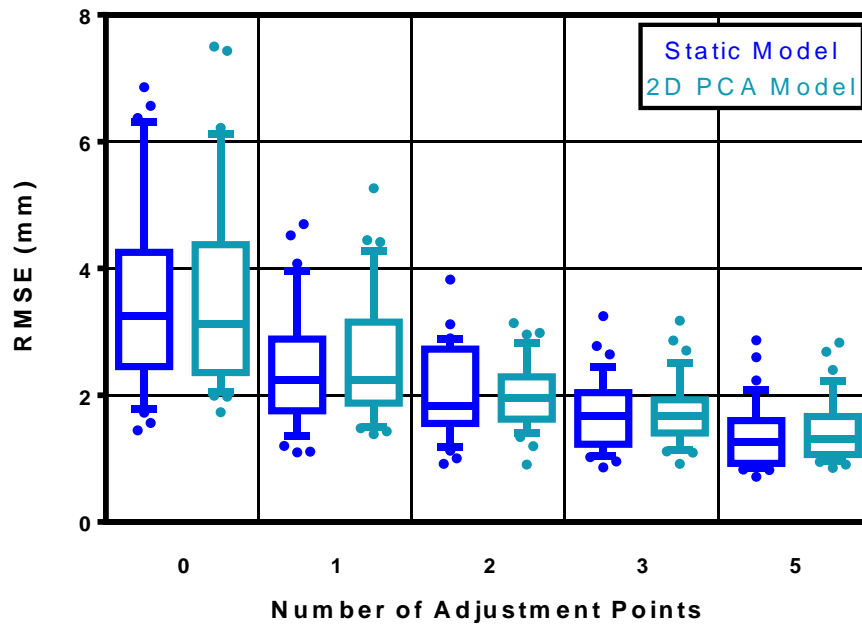


Figure 19: Forecast error of static and 2D PCA models – GTV centroid position descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

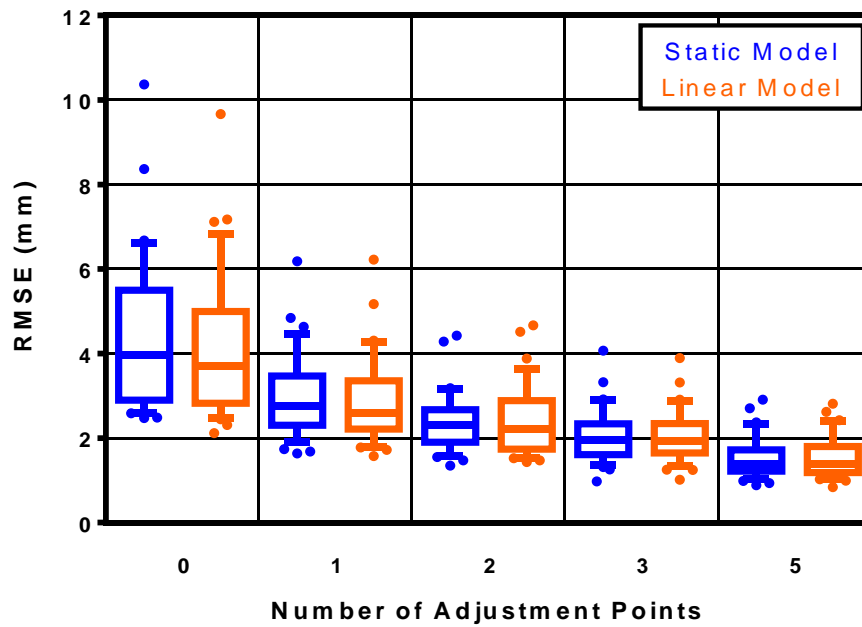


Figure 20: Forecast error of static and linear models – Radial extent GTV morphology descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

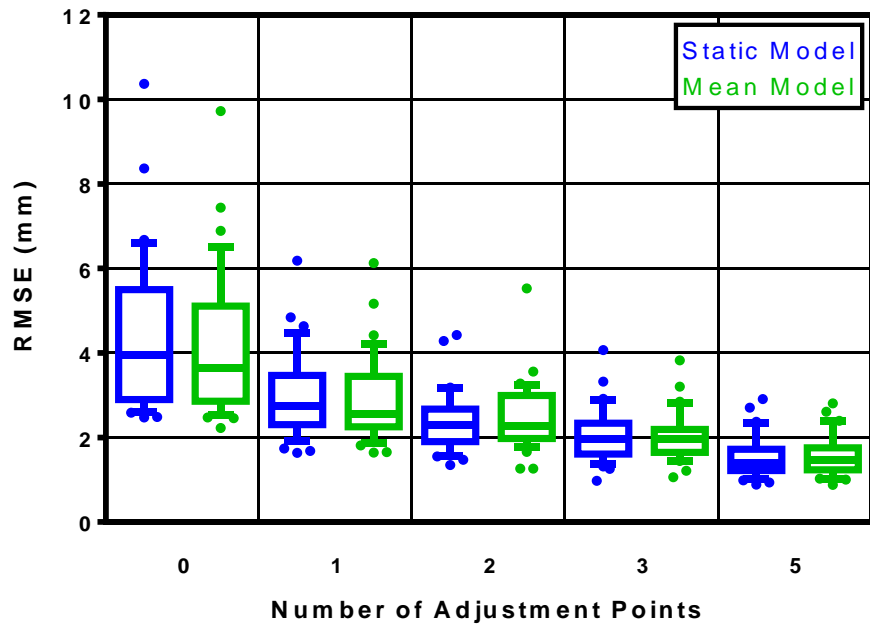


Figure 21: Forecast error of static and mean models – Radial extent GTV morphology descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

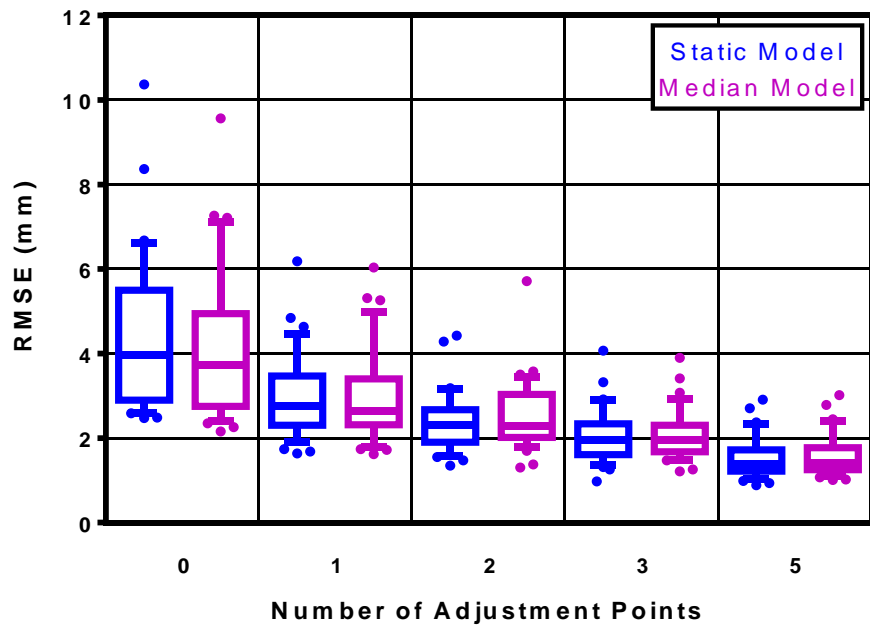


Figure 22: Forecast error of static and median models – Radial extent GTV morphology descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

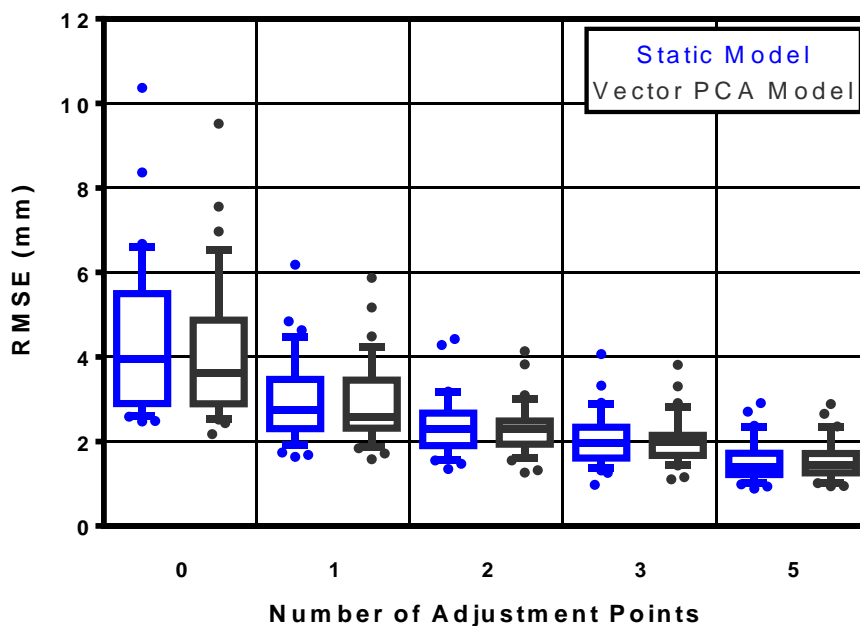


Figure 23: Forecast error of static and vector PCA models – Radial extent GTV morphology descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

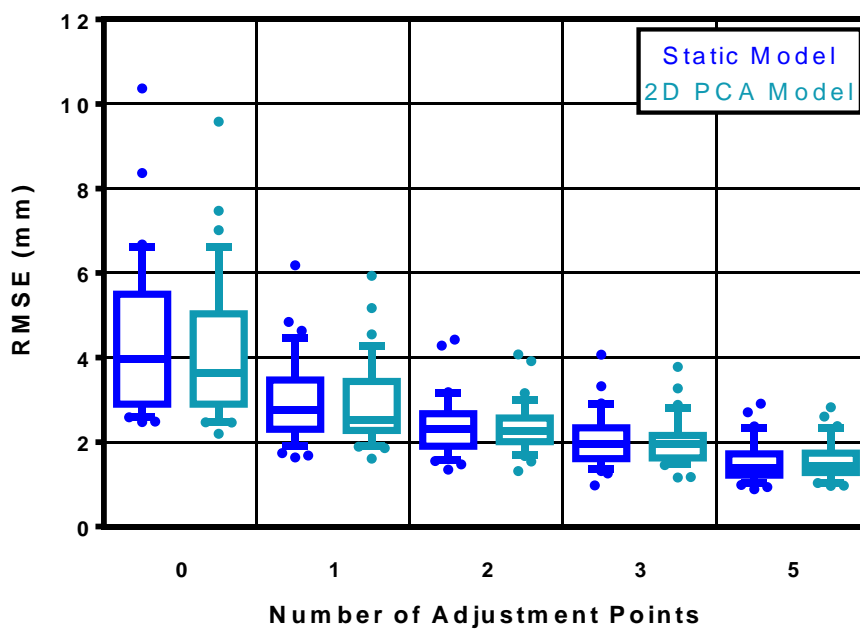


Figure 24: Forecast error of static and 2D PCA models – Radial extent GTV morphology descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

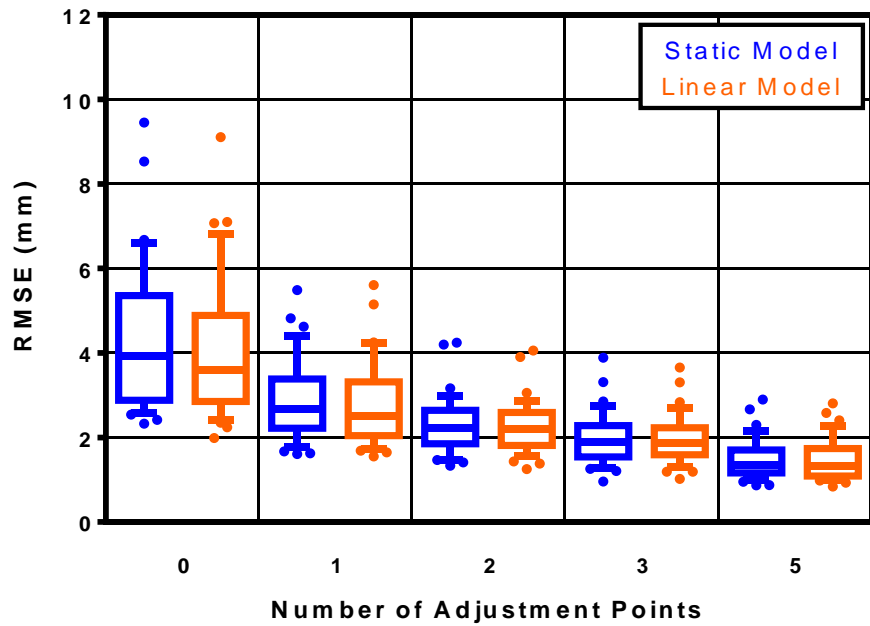


Figure 25: Forecast error of static and linear models – Spherical harmonic GTV morphology descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

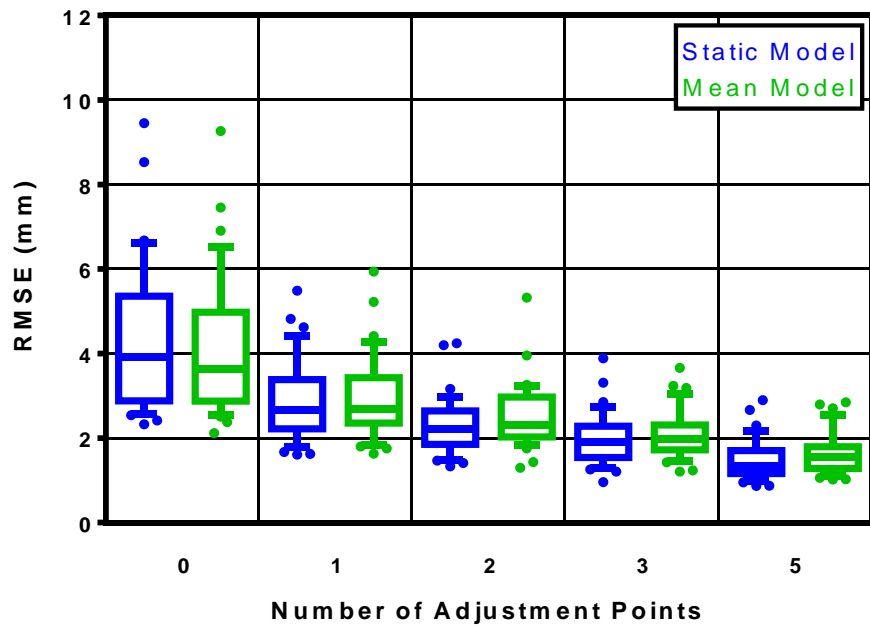


Figure 26: Forecast error of static and mean models – Spherical harmonic GTV morphology descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

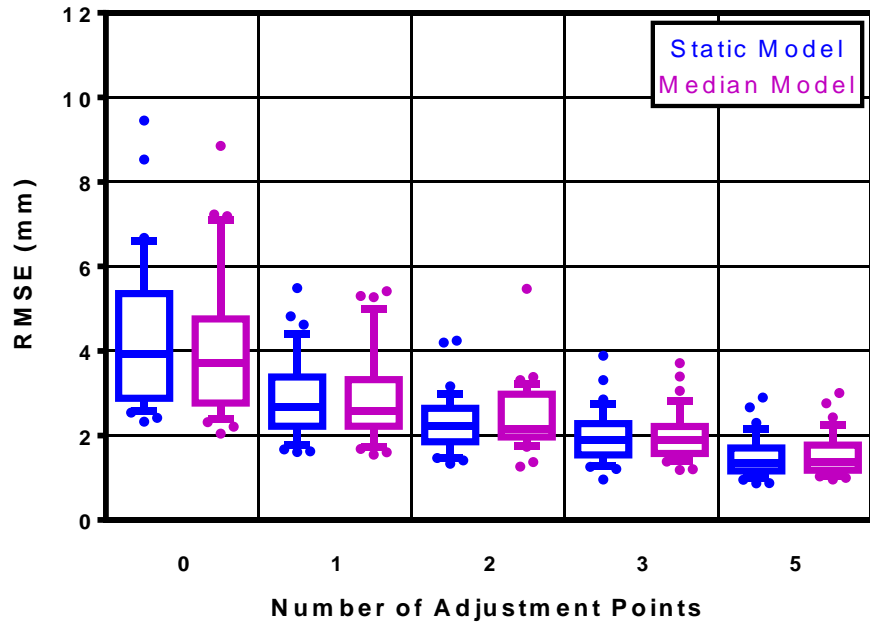


Figure 27: Forecast error of static and median models – Spherical harmonic GTV morphology descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

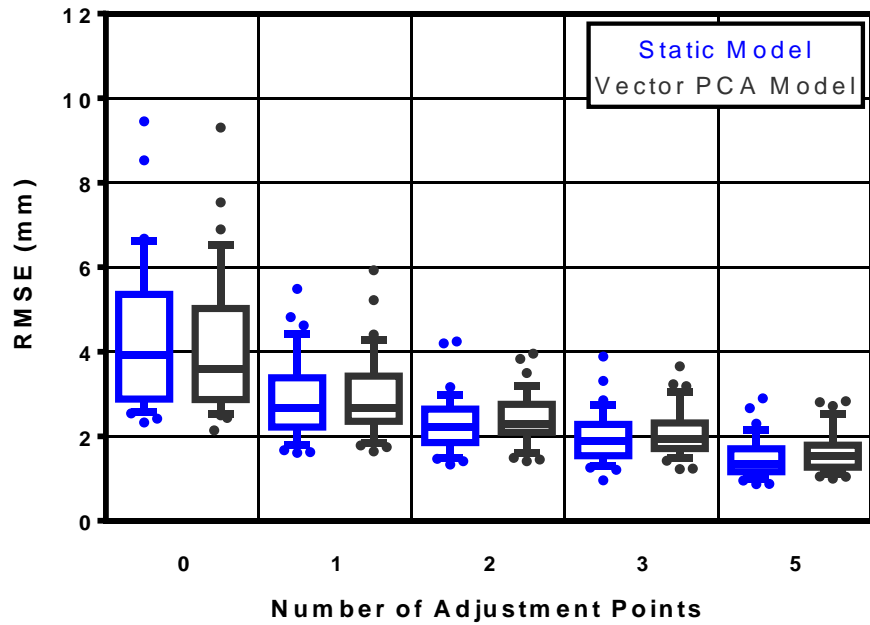


Figure 28: Forecast error of static and vector PCA models – Spherical harmonic GTV morphology descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

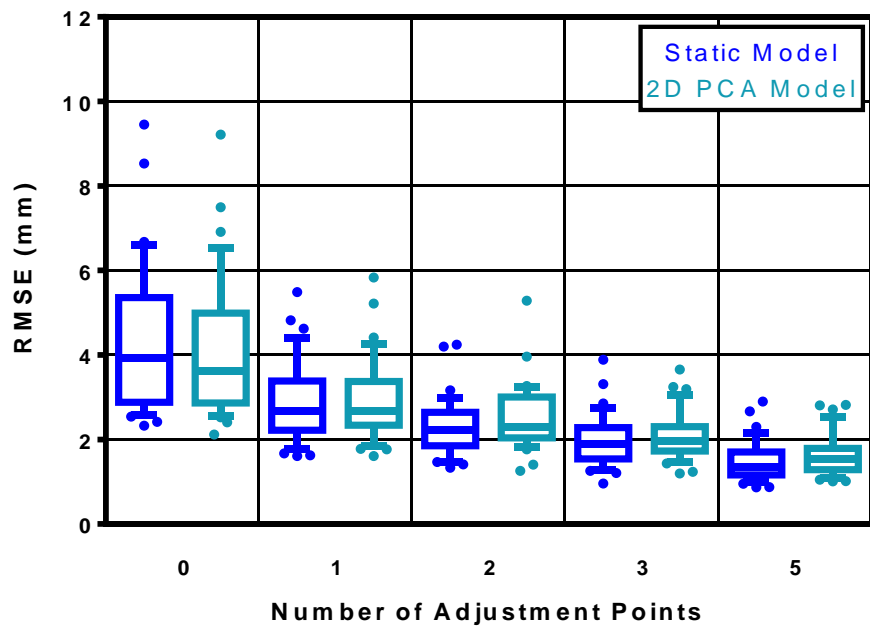


Figure 29: Forecast error of static and 2D PCA models – Spherical harmonic GTV morphology descriptor. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

4.2.1. Specific Aim 1 results

For Specific Aim 1, models of GTV volume and GTV centroid position were compared with other models of the same type but different numbers of adjustment points.

4.2.1.1. Specific Aim 1 results – GTV volume descriptor

Figures 30 – 35 depict the change in forecast error when adjustment points are included for all models of GTV volume. Inclusion of any number of adjustment points resulted in a statistically significant decrease in forecast error. Even a single adjustment point was sufficient to decrease the error by 1.7% – 6.5% depending on the model type. This range represented a 12.9% – 30.0% improvement relative to the median forecast error from models of the same type but with no adjustment points.

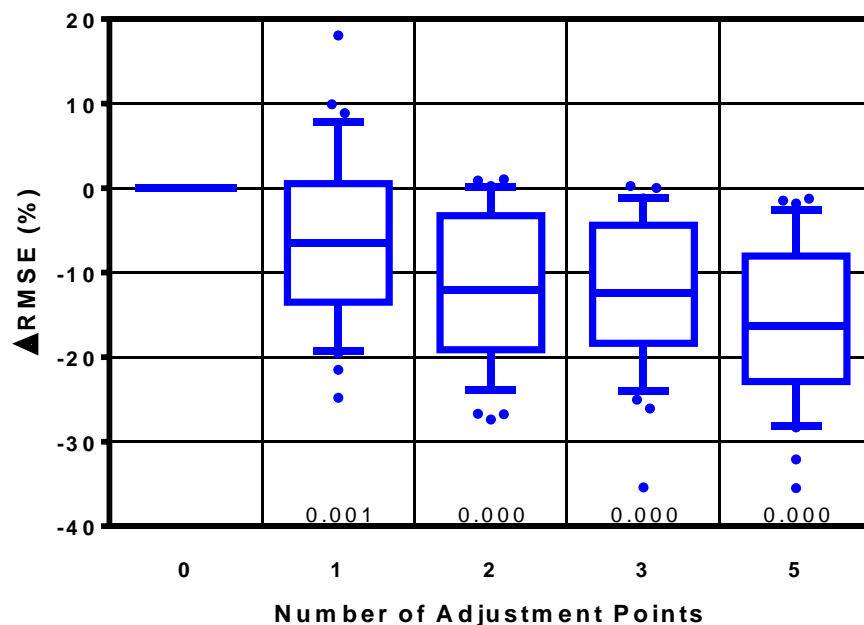


Figure 30: Change in forecast error of the static model with adjustment points – GTV volume descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

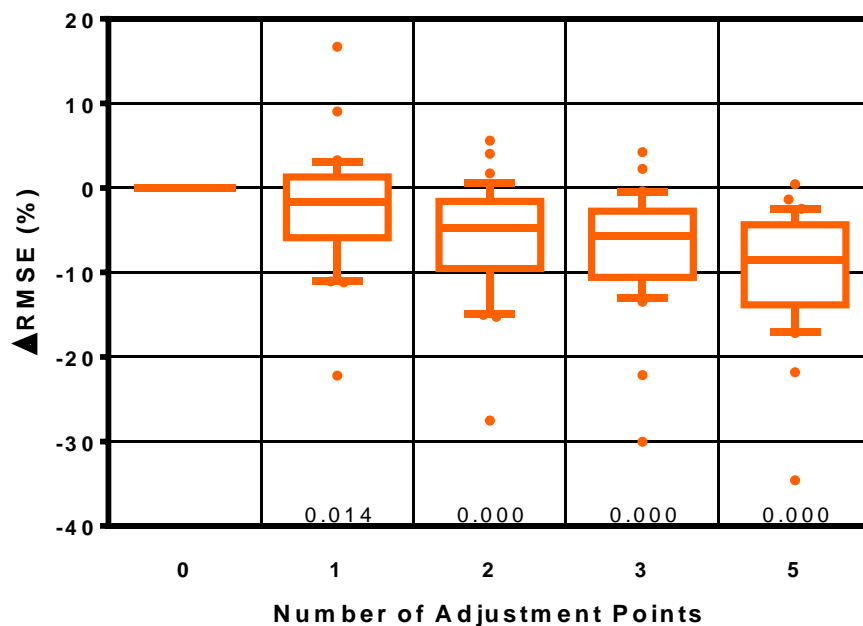


Figure 31: Change in forecast error of the linear model with adjustment points – GTV volume descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

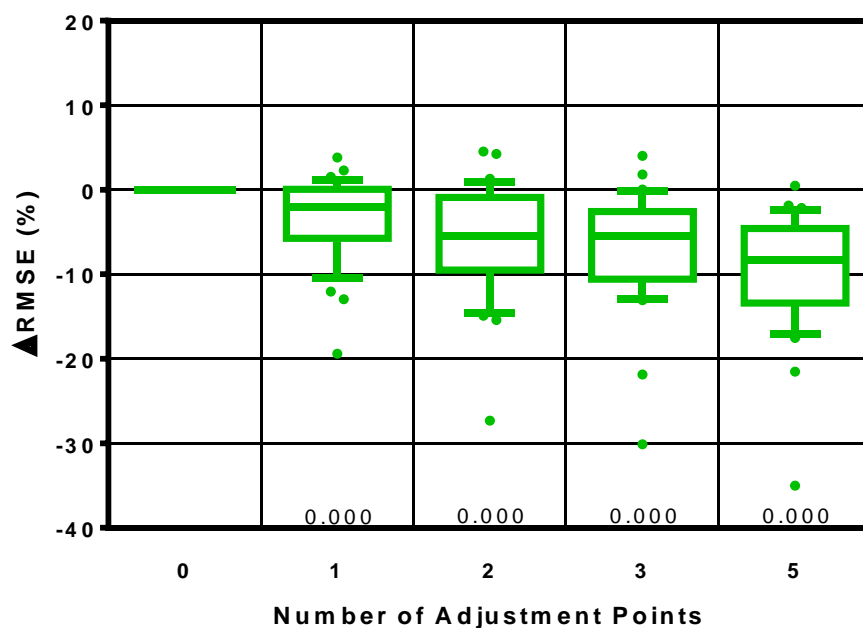


Figure 32: Change in forecast error of the mean model with adjustment points – GTV volume descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

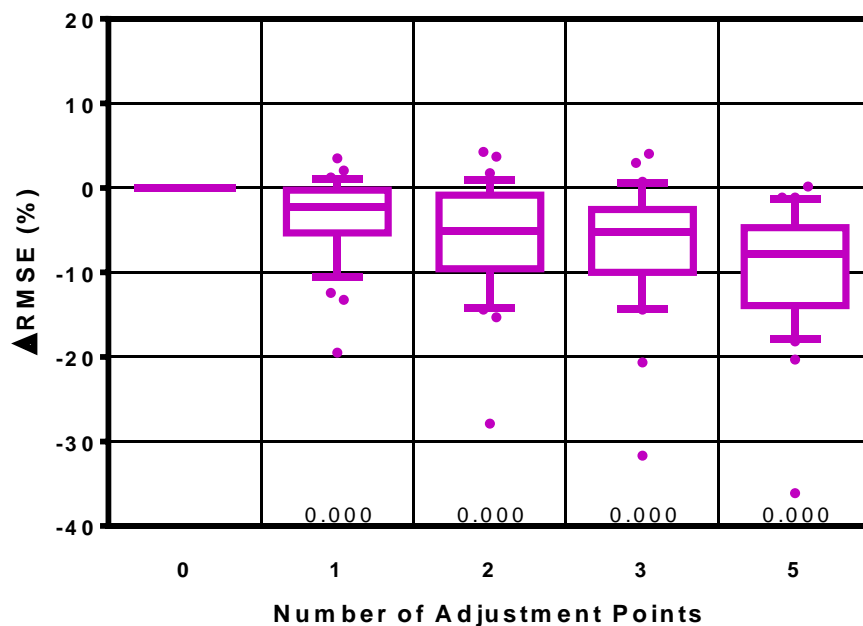


Figure 33: Change in forecast error of the median model with adjustment points – GTV volume descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

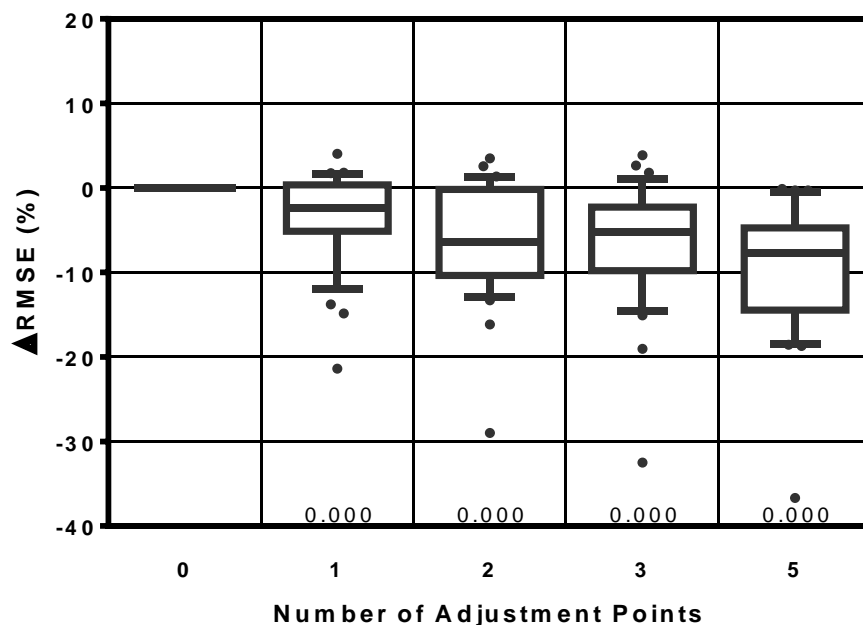


Figure 34: Change in forecast error of the vector PCA model with adjustment points – GTV volume descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

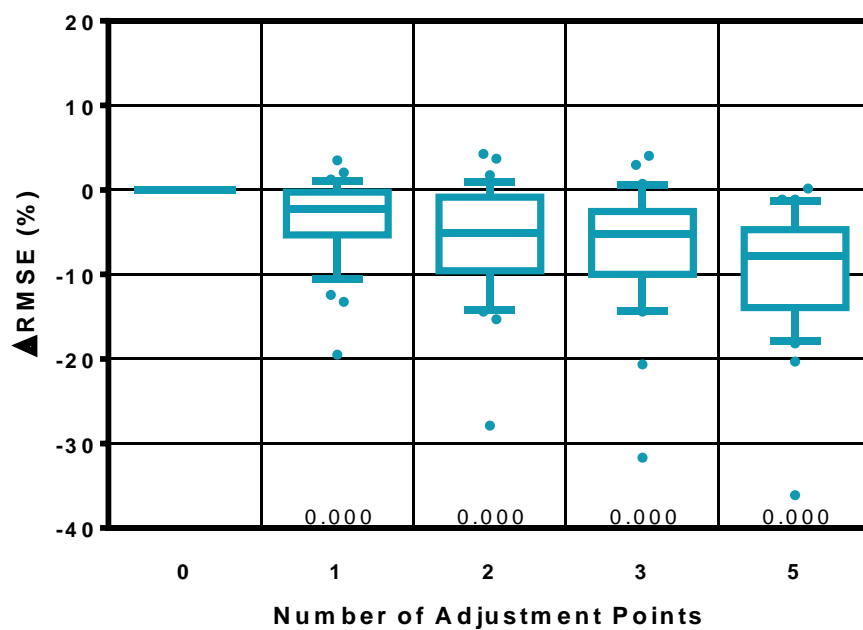


Figure 35: Change in forecast error of the 2D PCA model with adjustment points – GTV volume descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

Figures 36 – 41 present the sequential improvement in forecast accuracy with each subsequent addition of an adjustment point. For all model types, the additions of the first and second adjustment points were both statistically significant. However, the addition of the third was not, except for the static model. Adding two more (for a total of 5 adjustment points) again led to a statistically significant decrease in forecast error.

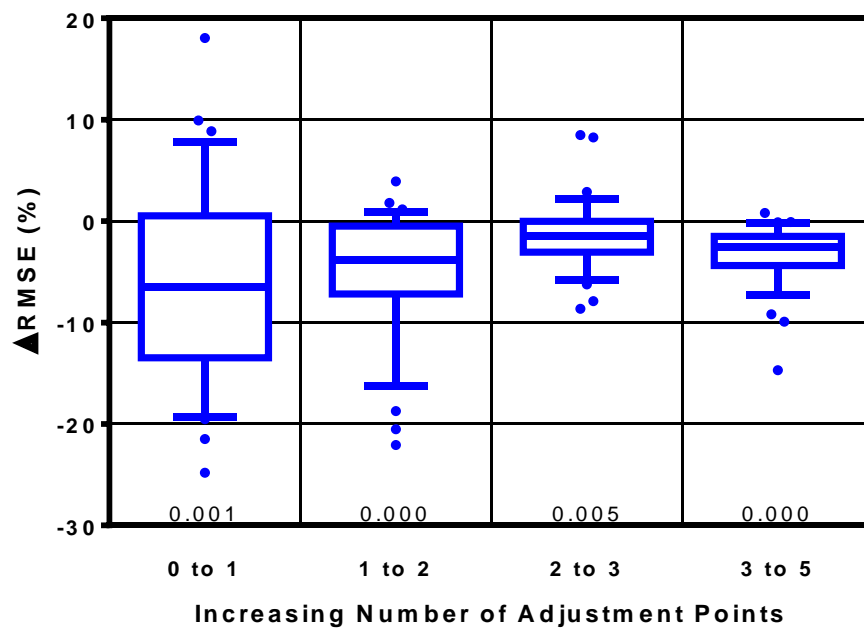


Figure 36: Marginal change in forecast error of the static model with subsequent adjustment points – GTV volume descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

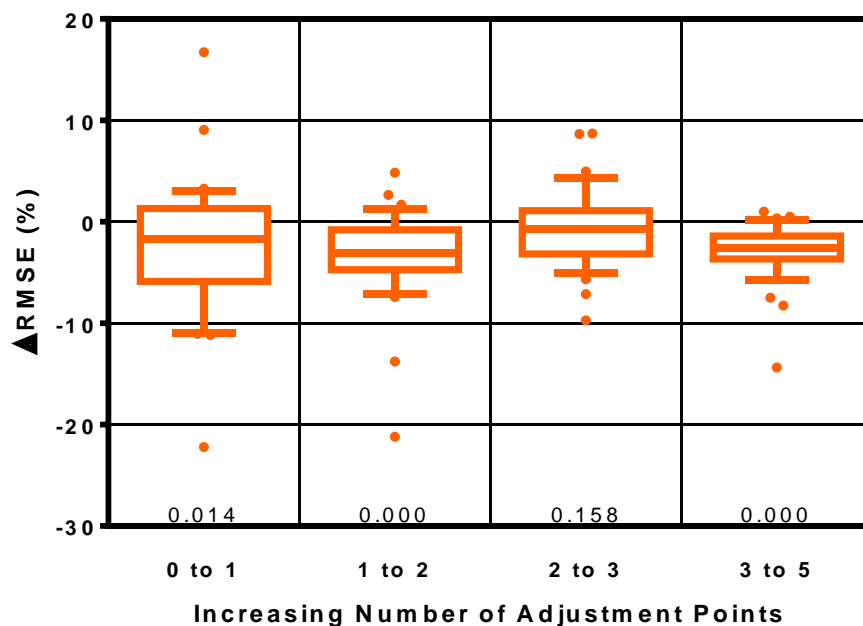


Figure 37: Marginal change in forecast error of the linear model with subsequent adjustment points – GTV volume descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

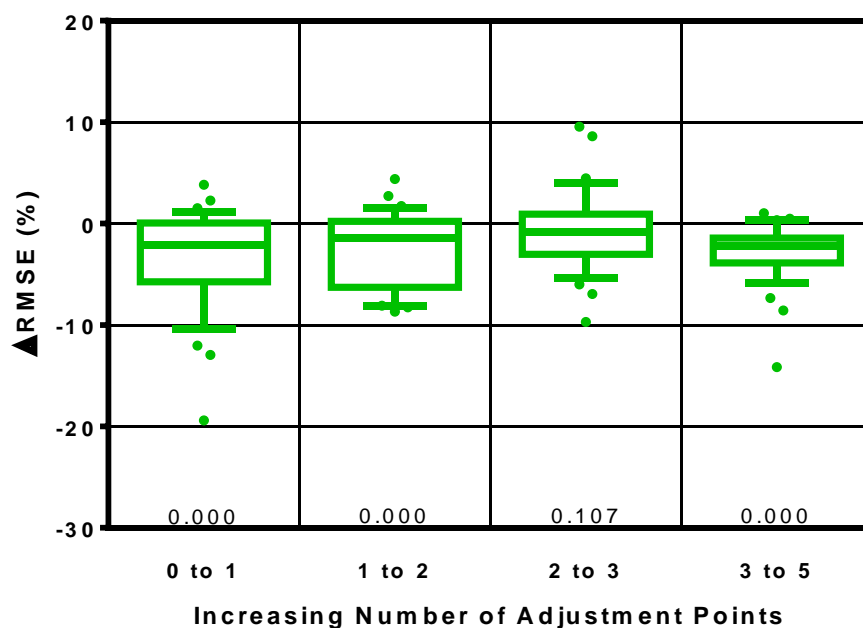


Figure 38: Marginal change in forecast error of the mean model with subsequent adjustment points – GTV volume descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

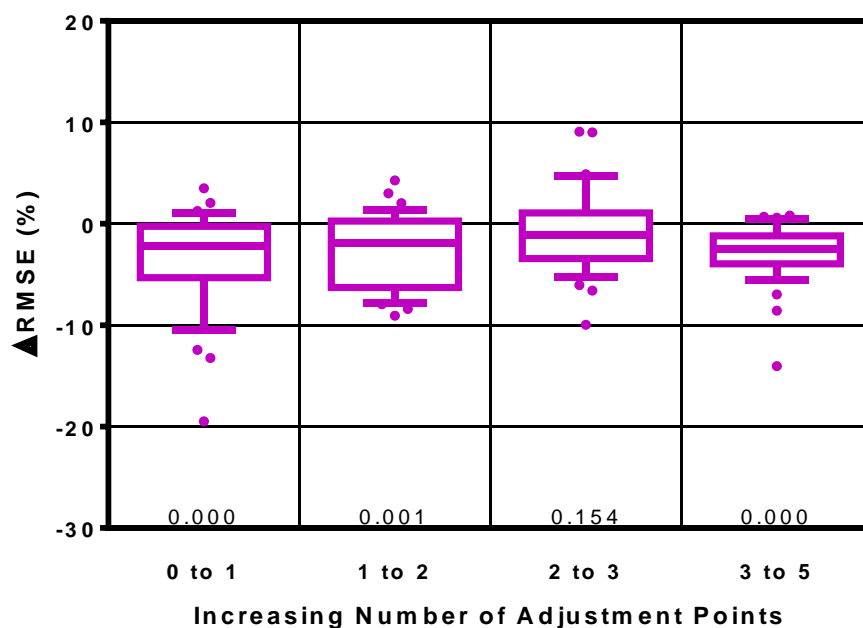


Figure 39: Marginal change in forecast error of the median model with subsequent adjustment points – GTV volume descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

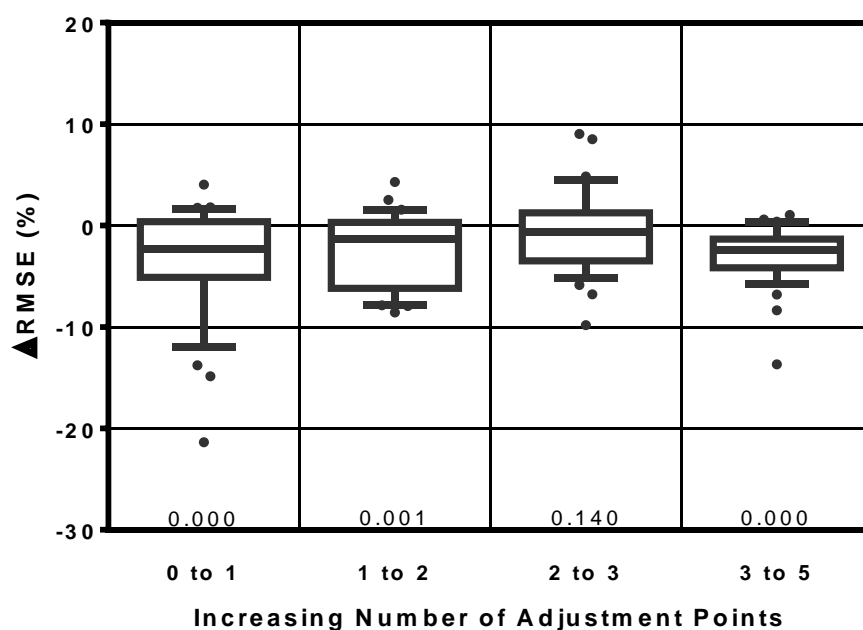


Figure 40: Marginal change in forecast error of the vector PCA model with subsequent adjustment points – GTV volume descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

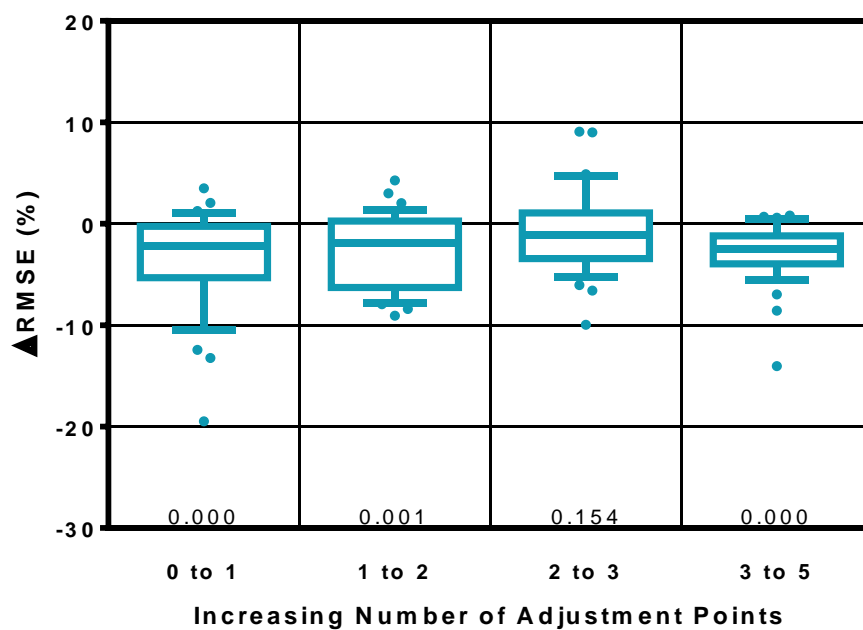


Figure 41: Marginal change in forecast error of the 2D PCA model with subsequent adjustment points – GTV volume descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

4.2.1.2. Specific Aim 1 results – GTV centroid position descriptor

The forecast error of models of the GTV centroid position depended on the number of adjustment points in a way similar to that of the GTV volume. For all model types, the inclusion of any number of adjustment points significantly decreased the error relative to models with no adjustment points (Figures 42 – 47). The median decrease in error due to the addition of a single adjustment point was 0.9 mm – 1.1 mm, depending on the model type. This represented 27.0% – 34.1% of the median forecast error of models of the same type but with no adjustment points.

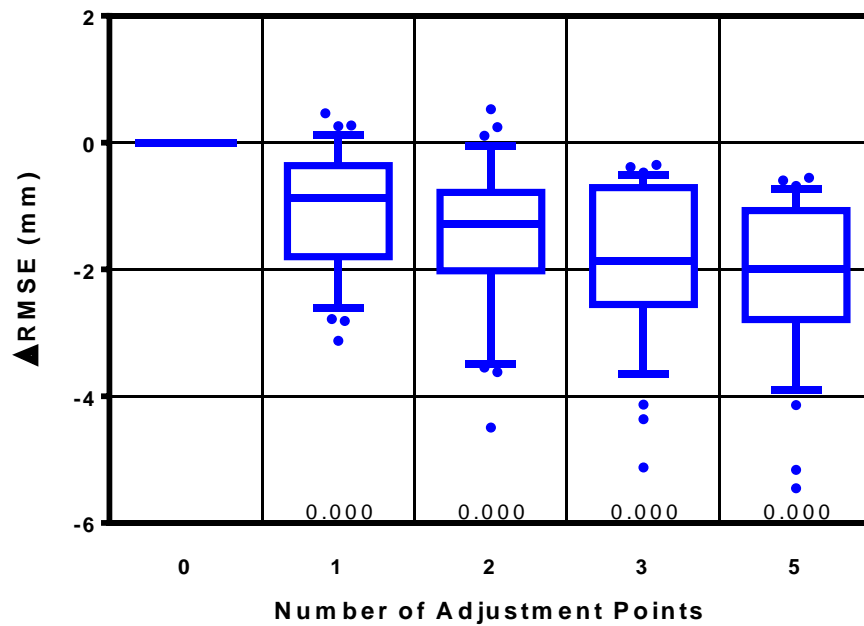


Figure 42: Change in forecast error of the static model with adjustment points – GTV centroid position descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

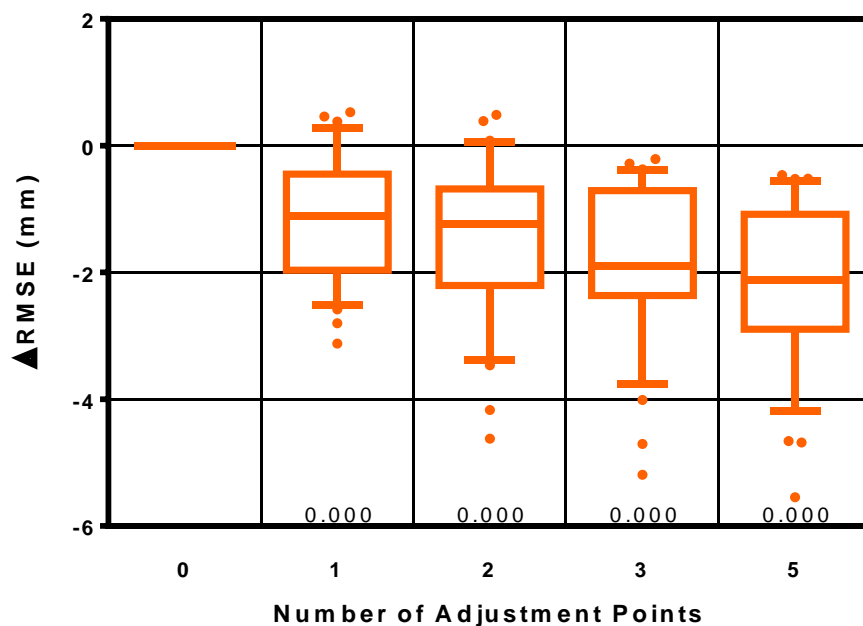


Figure 43: Change in forecast error of the linear model with adjustment points – GTV centroid position descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

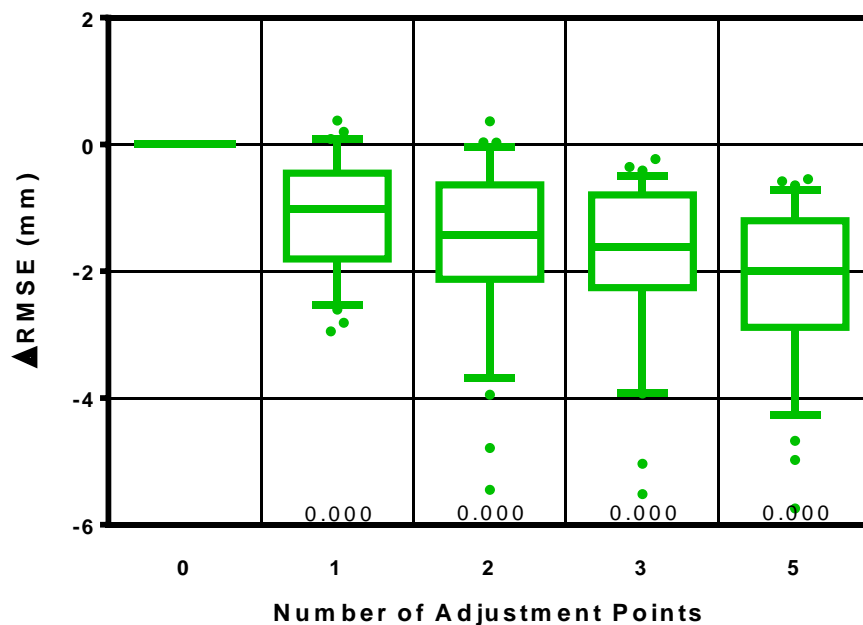


Figure 44: Change in forecast error of the mean model with adjustment points – GTV centroid position descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

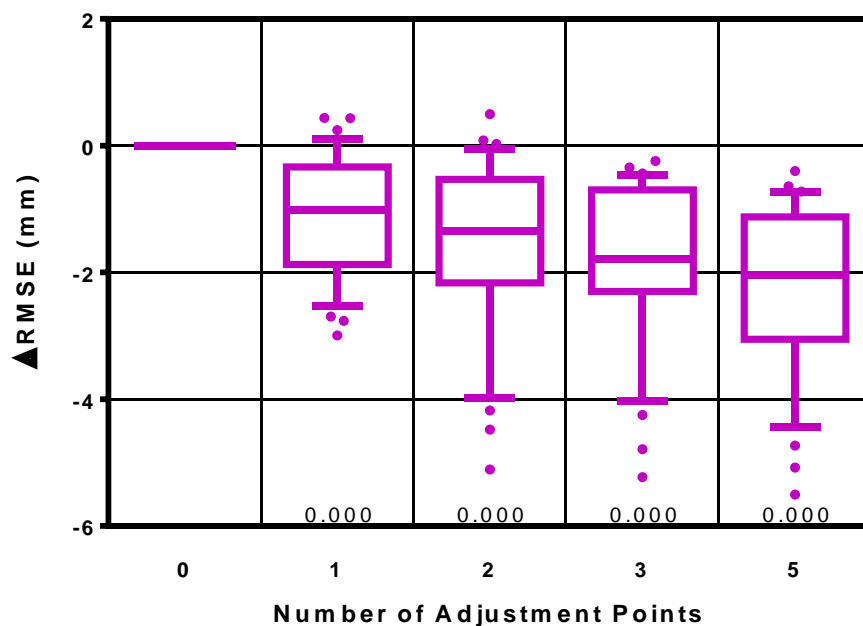


Figure 45: Change in forecast error of the median model with adjustment points – GTV centroid position descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

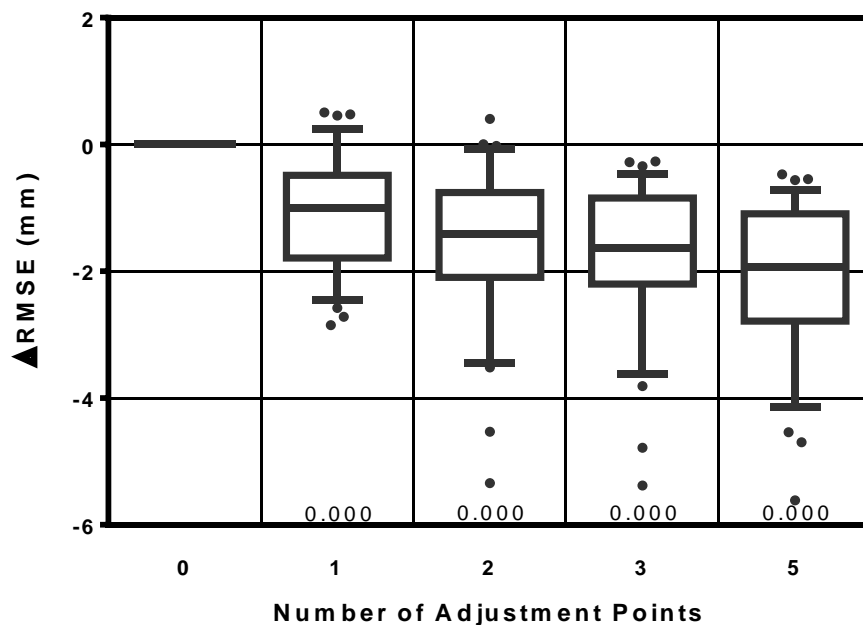


Figure 46: Change in forecast error of the vector PCA model with adjustment points – GTV centroid position descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

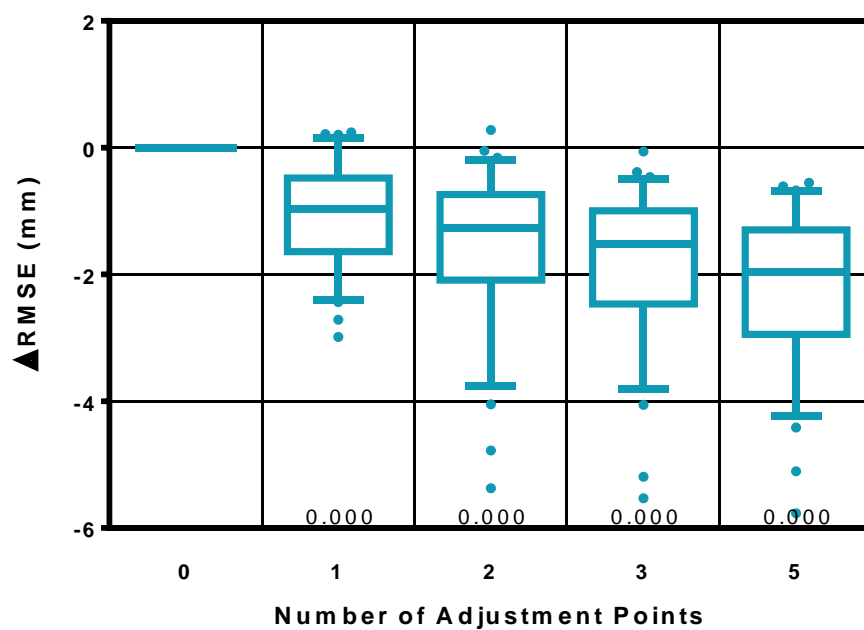


Figure 47: Change in forecast error of the 2D PCA model with adjustment points – GTV centroid position descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

In addition, each successive addition of an adjustment point led to a further decrease in error that was statistically significant. (Figures 48 – 53). The magnitude and range of the change in forecast error, however, decreased as the number of adjustment points increased.

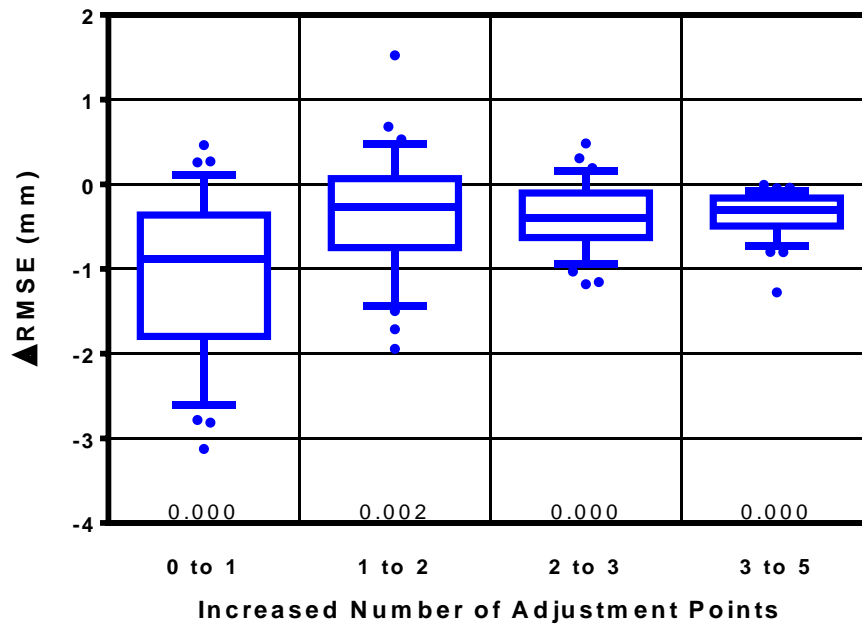


Figure 48: Marginal change in forecast error of the static model with subsequent adjustment points – GTV centroid position descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

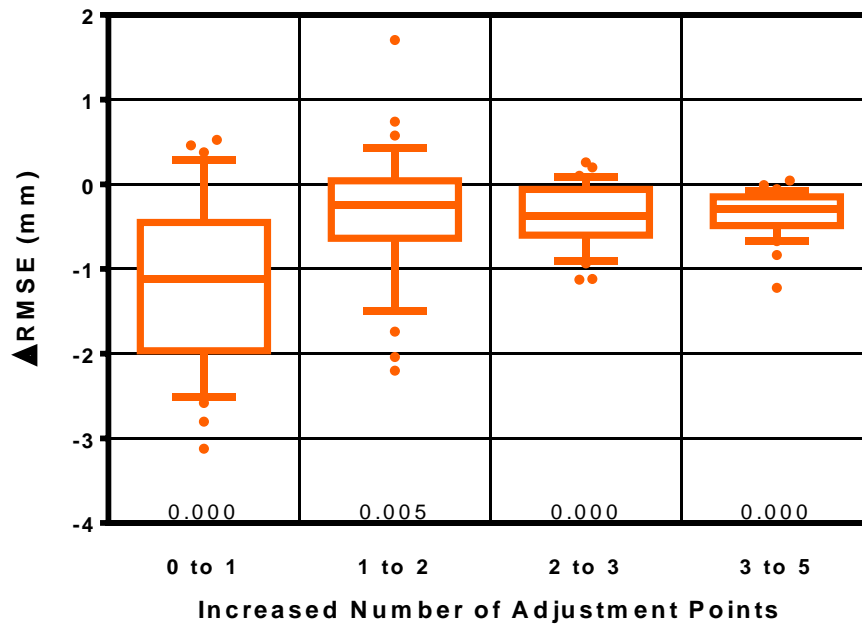


Figure 49: Marginal change in forecast error of the linear model with subsequent adjustment points – GTV centroid position descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

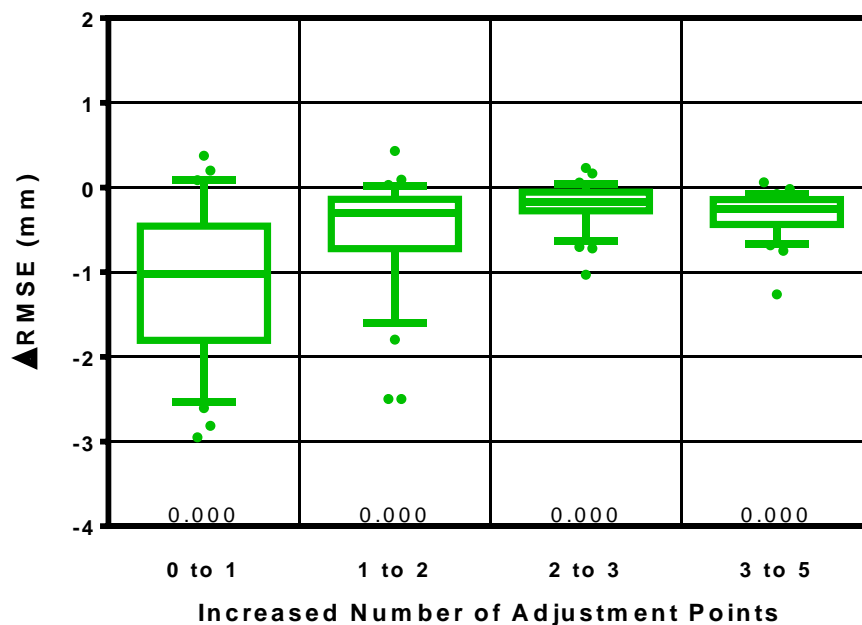


Figure 50: Marginal change in forecast error of the mean model with subsequent adjustment points – GTV centroid position descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

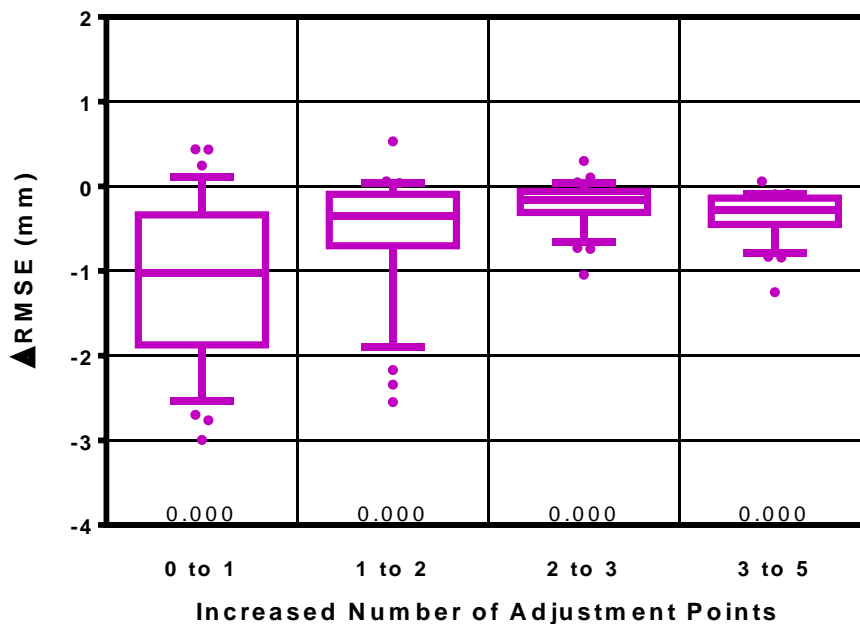


Figure 51: Marginal change in forecast error of the median model with subsequent adjustment points – GTV centroid position descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

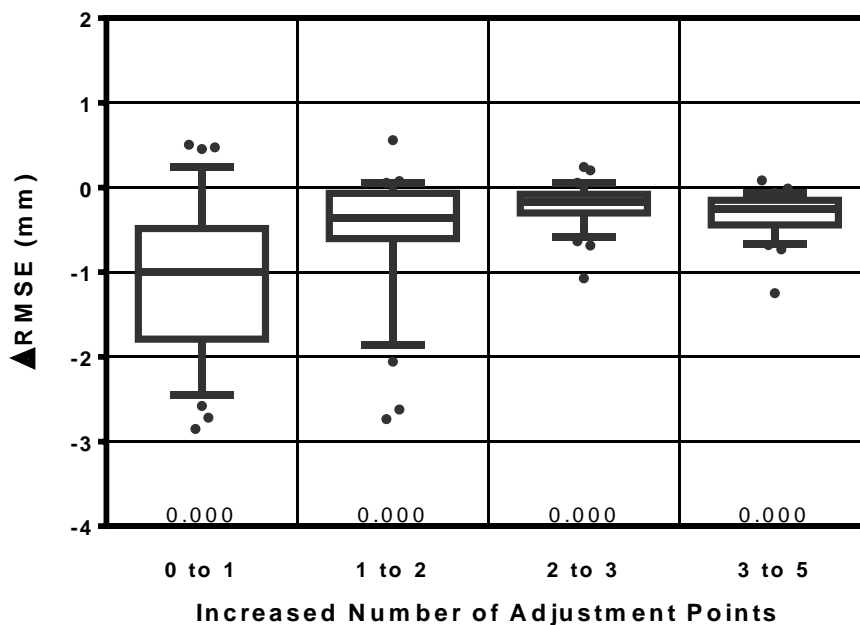


Figure 52: Marginal change in forecast error of the vector PCA model with subsequent adjustment points – GTV centroid position descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

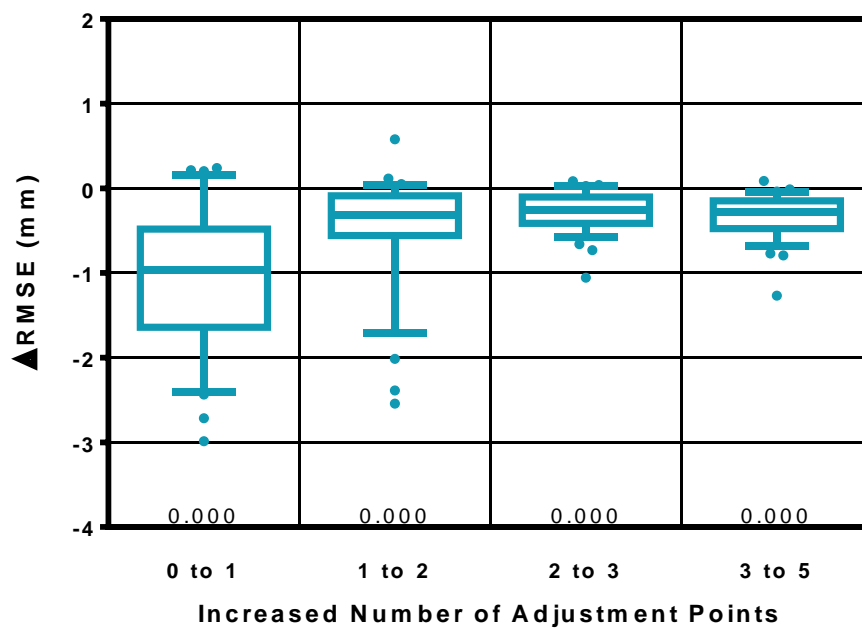


Figure 53: Marginal change in forecast error of the 2D PCA model with subsequent adjustment points – GTV centroid position descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

4.2.2. Specific Aim 2 results

As in Specific Aim 1, Specific Aim 2 compares the forecast error of models of the same type but with differing numbers of adjustment points. However, in Specific Aim 2, the models are based on the radial extent and spherical harmonic GTV morphology descriptors.

4.2.2.1. Specific Aim 2 results – Radial extent GTV morphology descriptor

Figures 54 – 59 depict the change in forecast error of models based on the radial extent GTV morphology descriptor when adjustment points are included. The inclusion of any number of adjustment points led to a statistically significant decrease in forecast error with a median change of 1.0 mm – 1.2 mm. This represented a 27.5% – 33.8% improvement in forecast accuracy of these models compared to models of the same type but with no adjustment points.

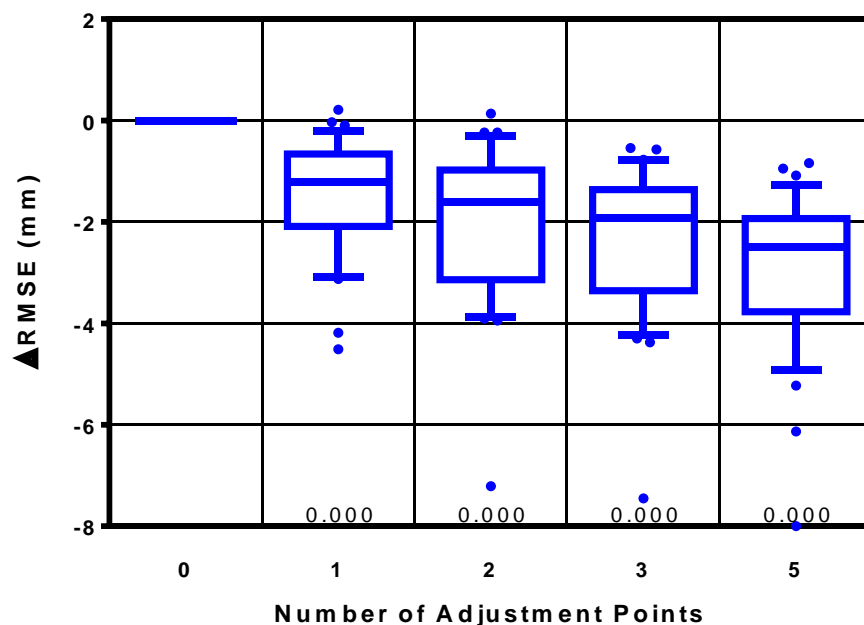


Figure 54: Change in forecast error of the static model with adjustment points – Radial extent GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

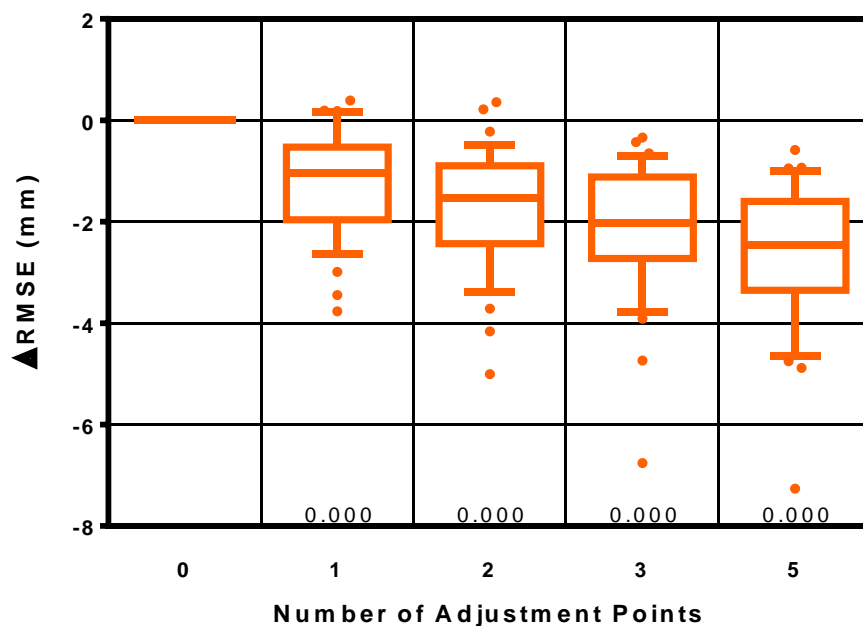


Figure 55: Change in forecast error of the linear model with adjustment points – Radial extent GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

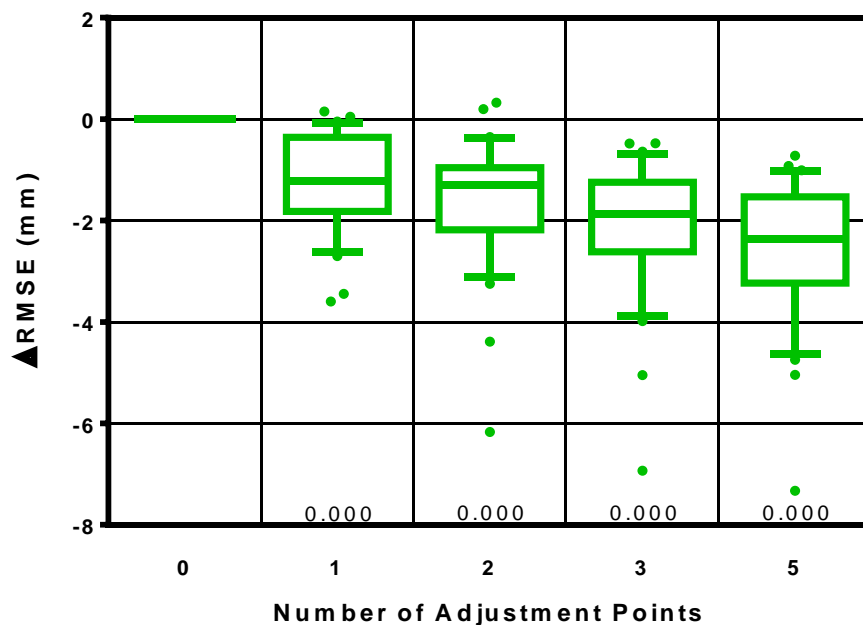


Figure 56: Change in forecast error of the mean model with adjustment points – Radial extent GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

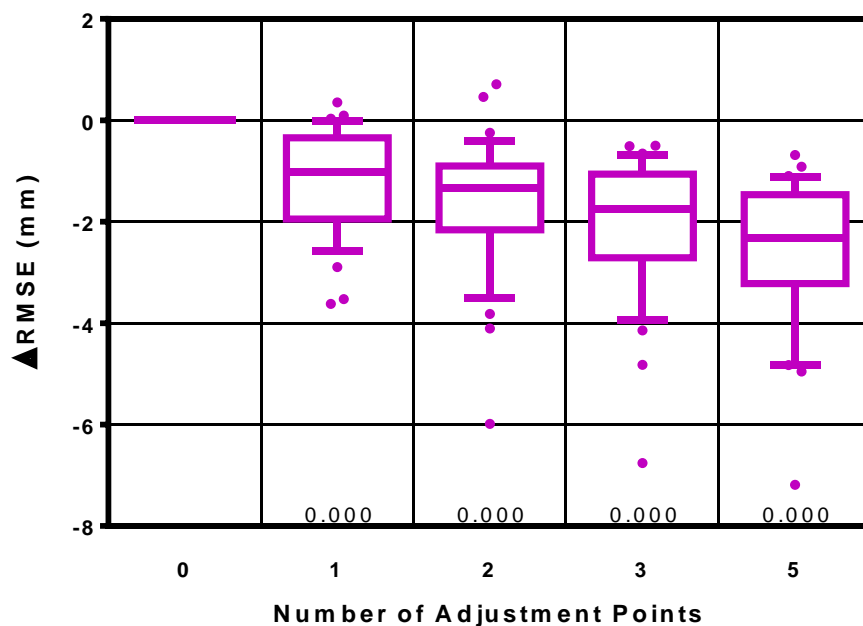


Figure 57: Change in forecast error of the median model with adjustment points – Radial extent GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

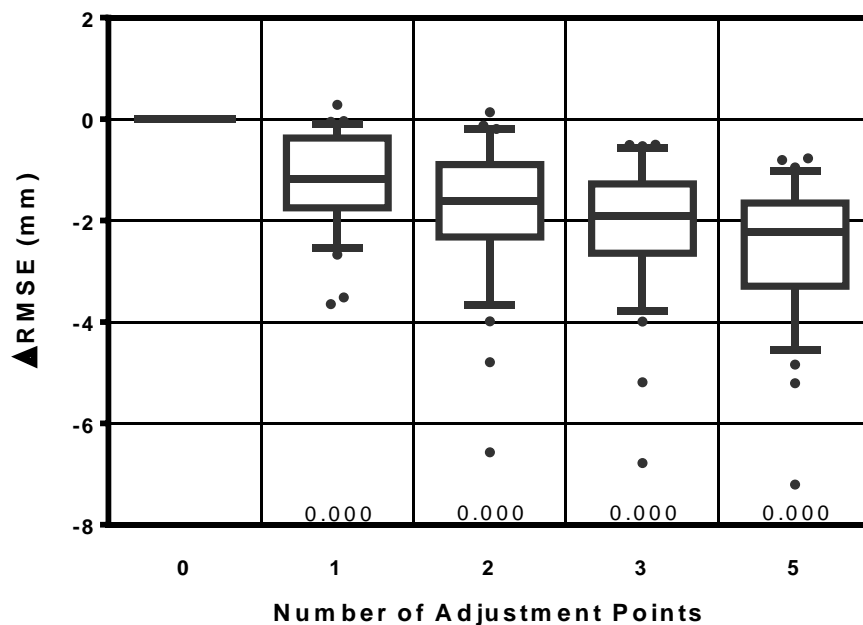


Figure 58: Change in forecast error of the vector PCA model with adjustment points – Radial extent GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

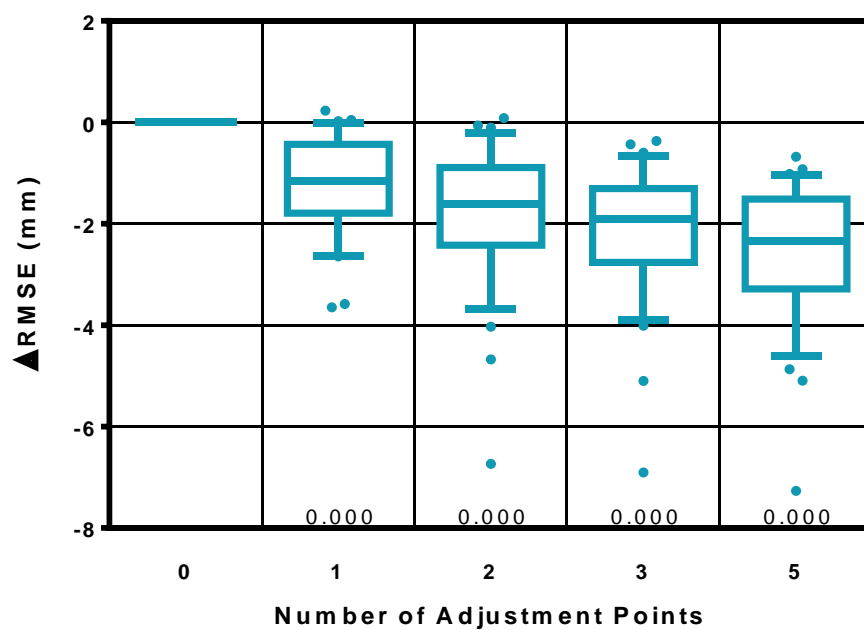


Figure 59: Change in forecast error of the 2D PCA model with adjustment points – Radial extent GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

Each successive addition of an adjustment point also led to a statistically significant decrease in error (Figures 60 – 65), although these changes decreased in magnitude and range as the number of adjustment points increased.

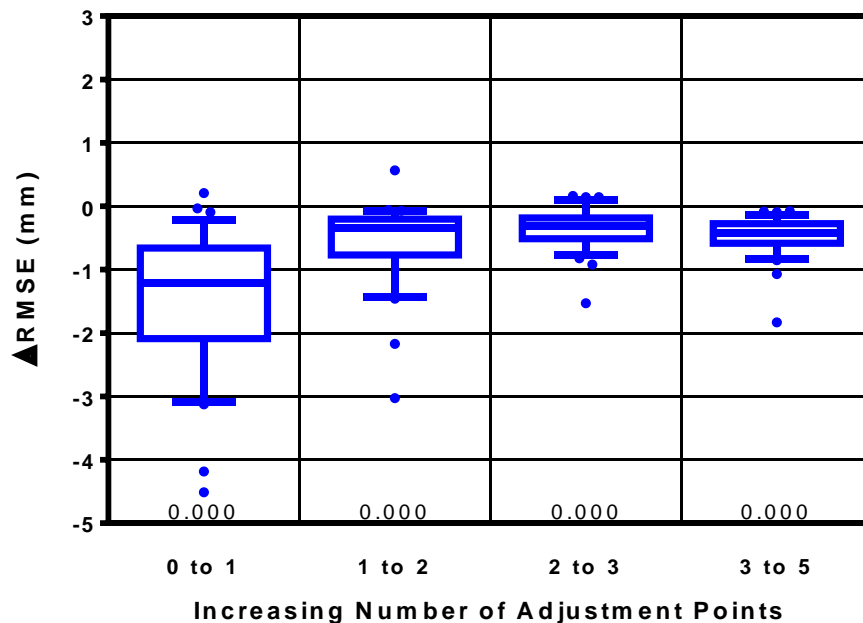


Figure 60: Marginal change in forecast error of the static model with subsequent adjustment points – Radial extent GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

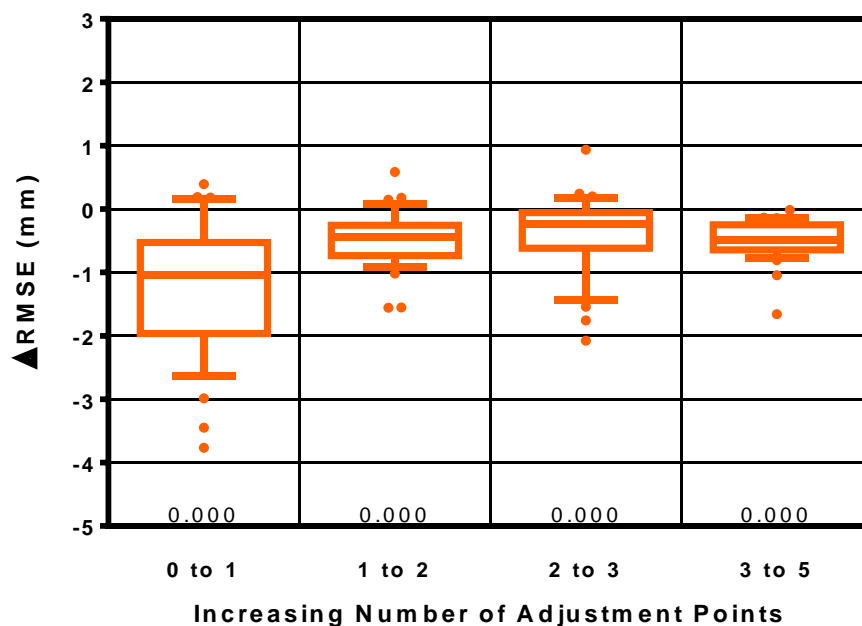


Figure 61: Marginal change in forecast error of the linear model with subsequent adjustment points – Radial extent GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

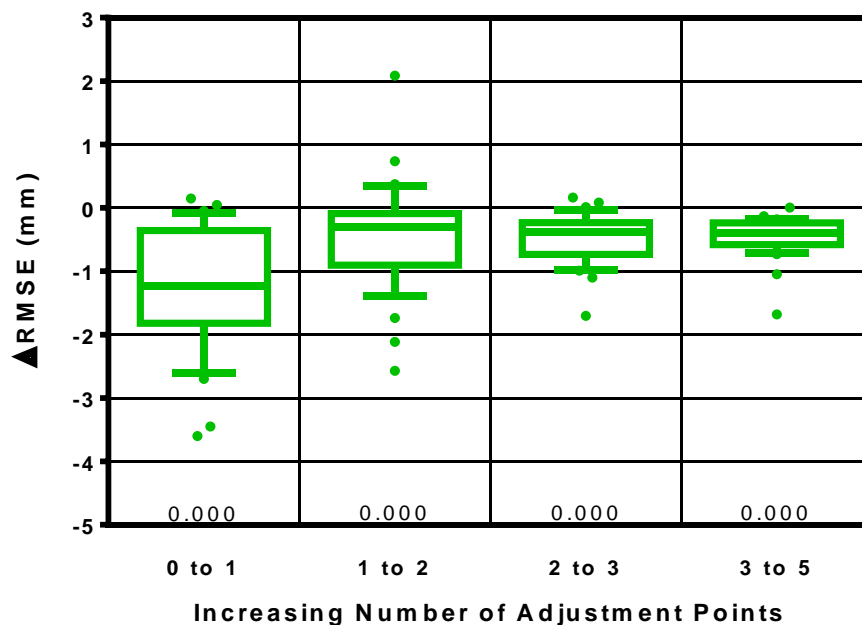


Figure 62: Marginal change in forecast error of the mean model with subsequent adjustment points – Radial extent GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

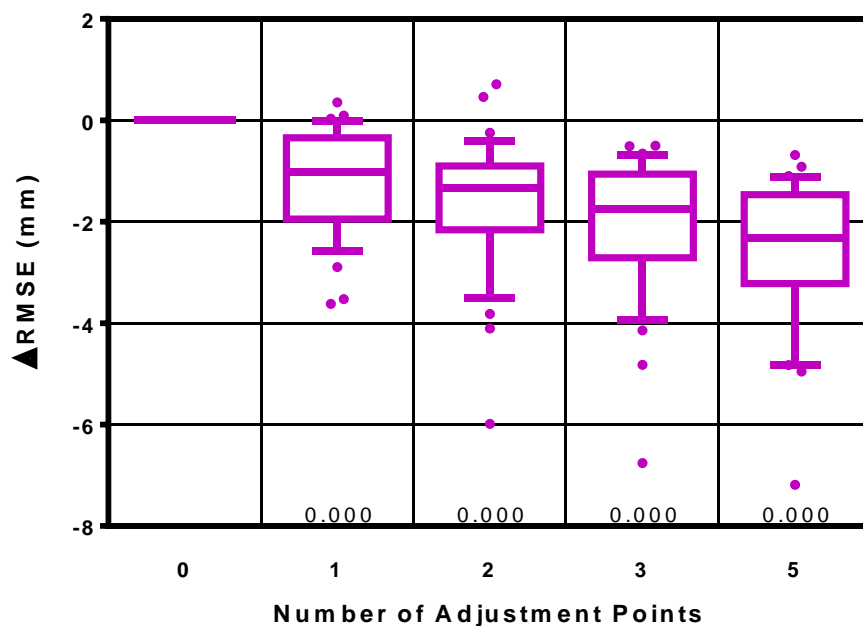


Figure 63: Marginal change in forecast error of the median model with subsequent adjustment points – Radial extent GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

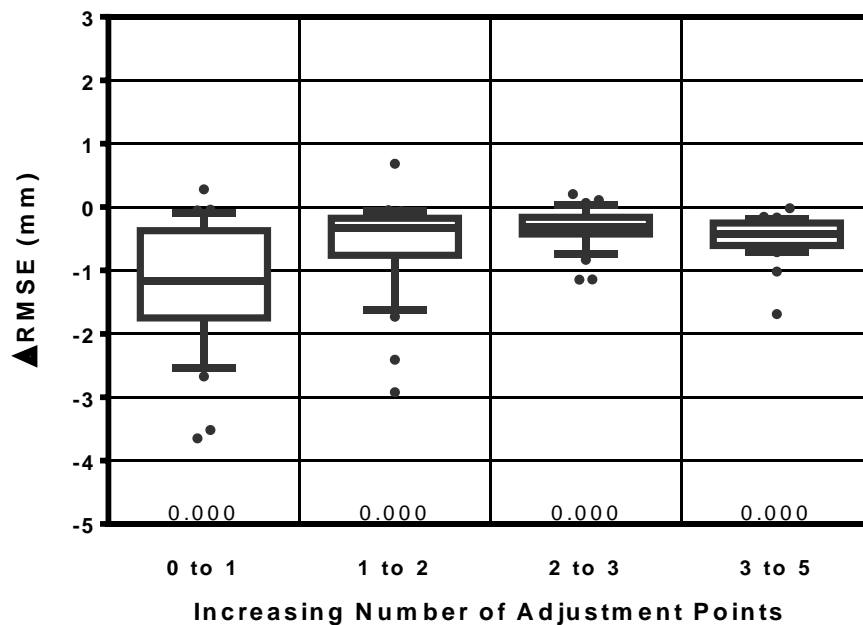


Figure 64: Marginal change in forecast error of the vector PCA model with subsequent adjustment points – Radial extent GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

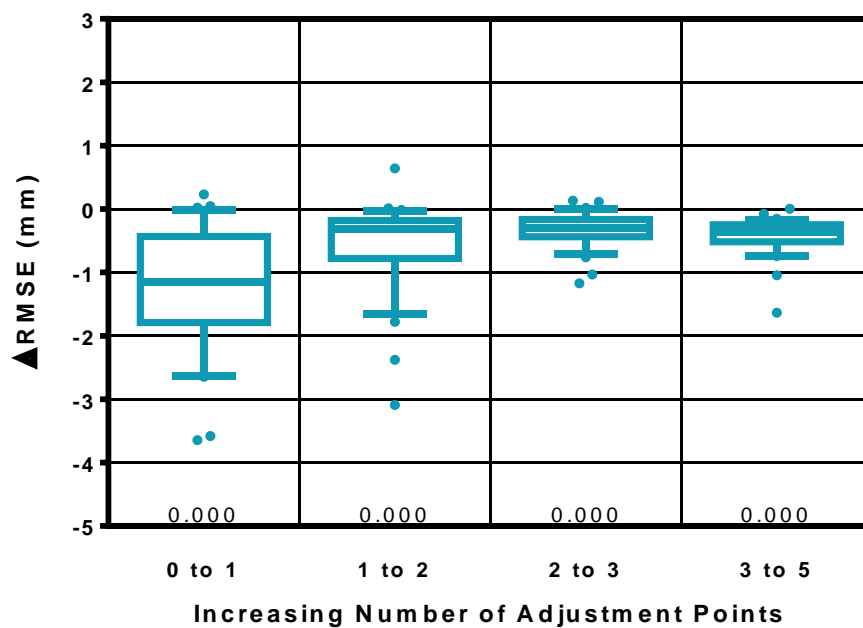


Figure 65: Marginal change in forecast error of the 2D PCA model with subsequent adjustment points – Radial extent GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

4.2.2.2. Specific Aim 2 results – Spherical harmonic GTV morphology descriptor

Models of the spherical harmonic GTV morphology descriptor depended on the number of adjustment points in a manner similar to that of the radial extent descriptor. Statistically significant decreases in the forecast error of models with any number of adjustment points compared to those with no adjustment points are depicted in Figures 66 – 71. The median change in error of adding the first adjustment point was 1.1 mm – 1.2 mm, depending on the model type. These changes represented 28.6% – 33.0% of the median error of models of the same type but with no adjustment points.

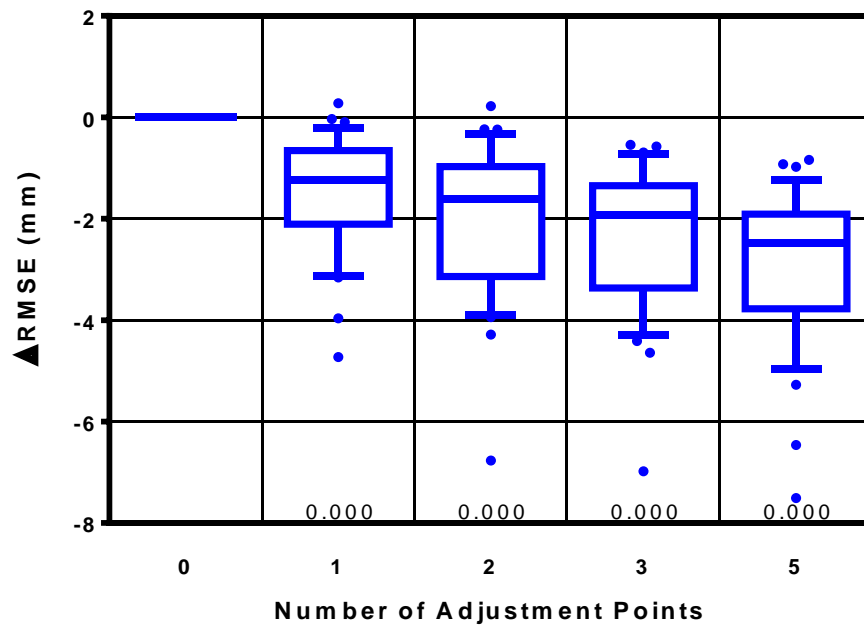


Figure 66: Change in forecast error of the static model with adjustment points – Spherical harmonic GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

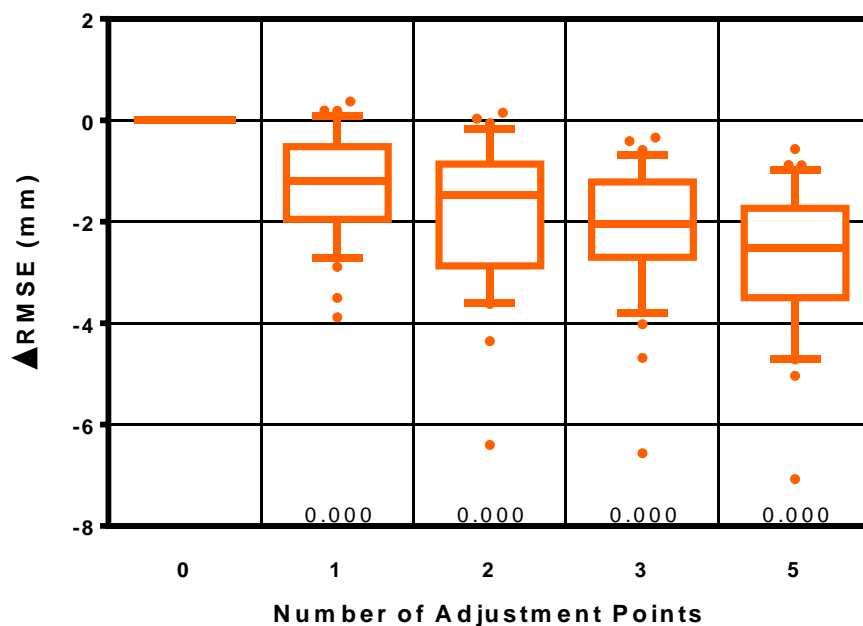


Figure 67: Change in forecast error of the linear model with adjustment points – Spherical harmonic GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

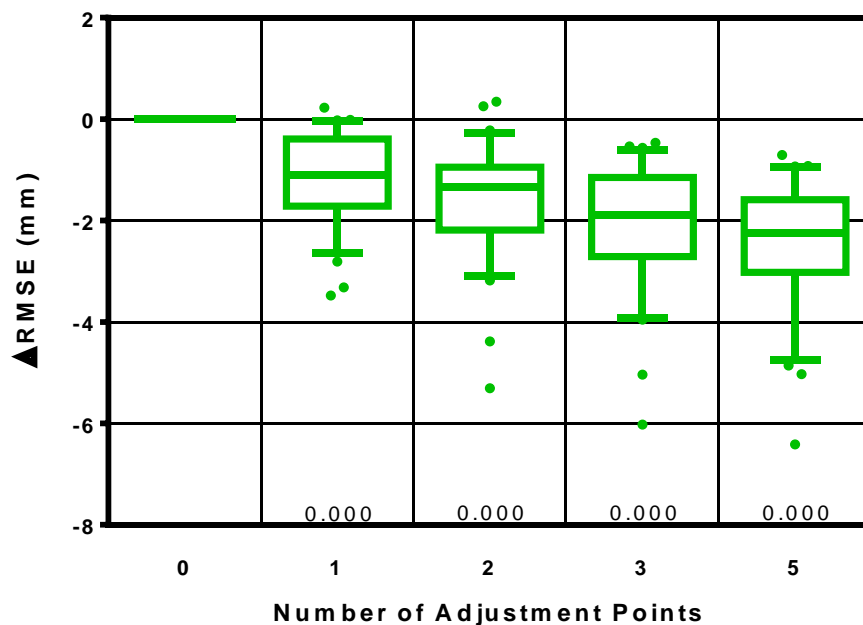


Figure 68: Change in forecast error of the mean model with adjustment points – Spherical harmonic GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

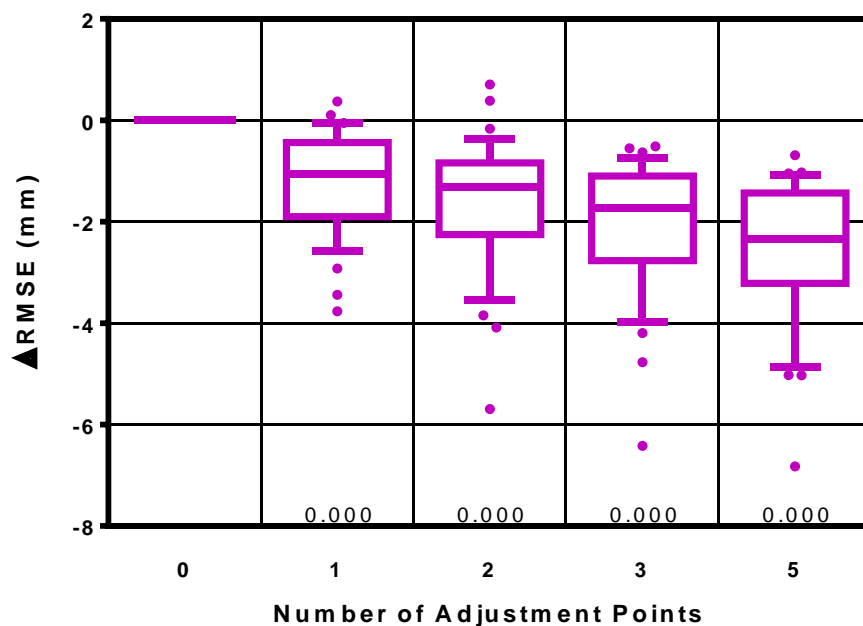


Figure 69: Change in forecast error of the median model with adjustment points – Spherical harmonic GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

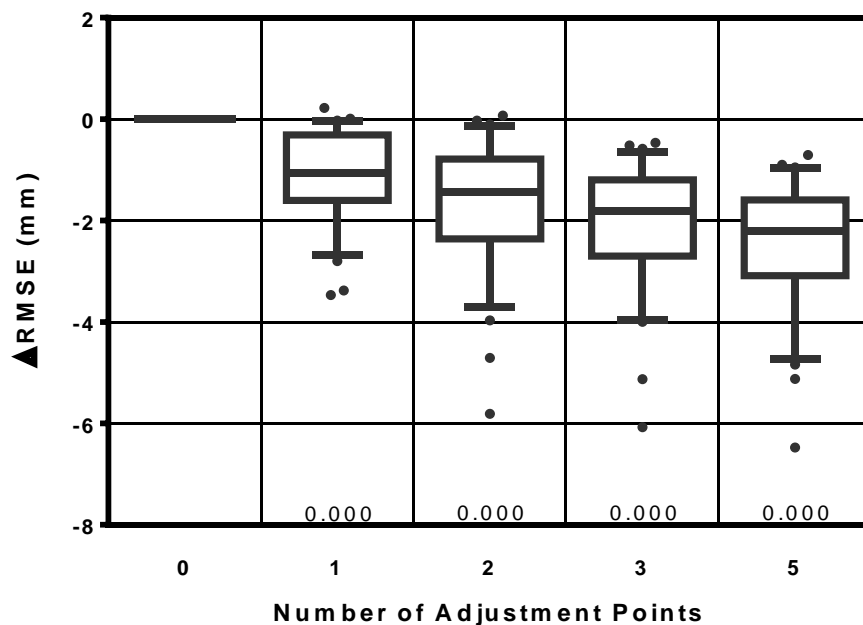


Figure 70: Change in forecast error of the vector PCA model with adjustment points – Spherical harmonic GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

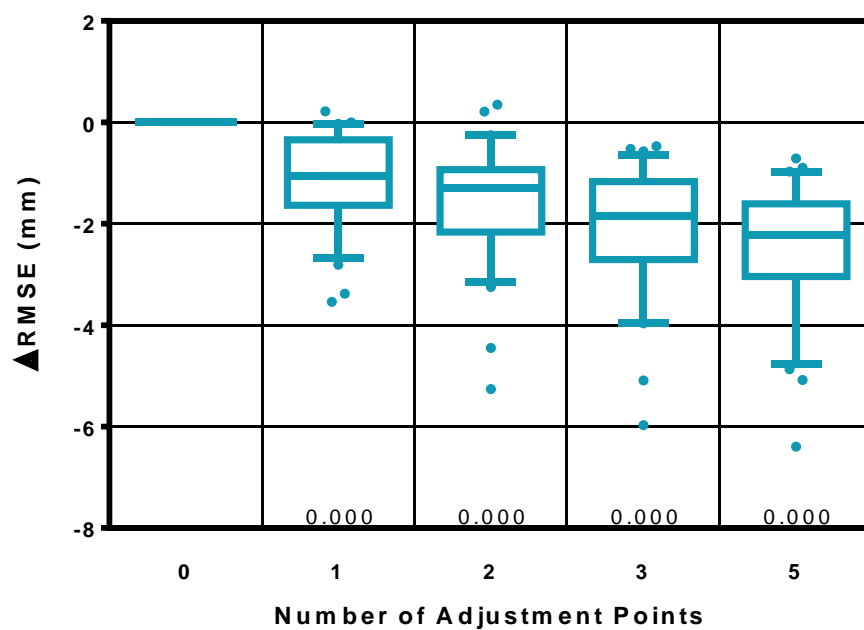


Figure 71: Change in forecast error of the 2D PCA model with adjustment points – Spherical harmonic GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

Each successive addition of an adjustment point resulted in a statistically significant decrease in the forecast error (Figures 72 – 77). As observed for the radial extent descriptor, the magnitude and range of these changes decreased with an increase in the number of adjustment points.

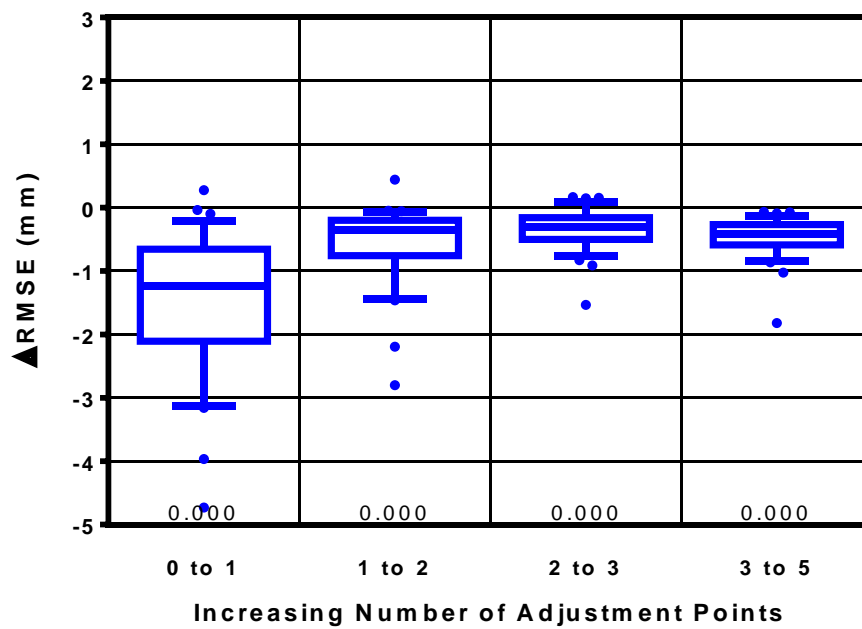


Figure 72: Marginal change in forecast error of the static model with subsequent adjustment points – Spherical harmonic GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

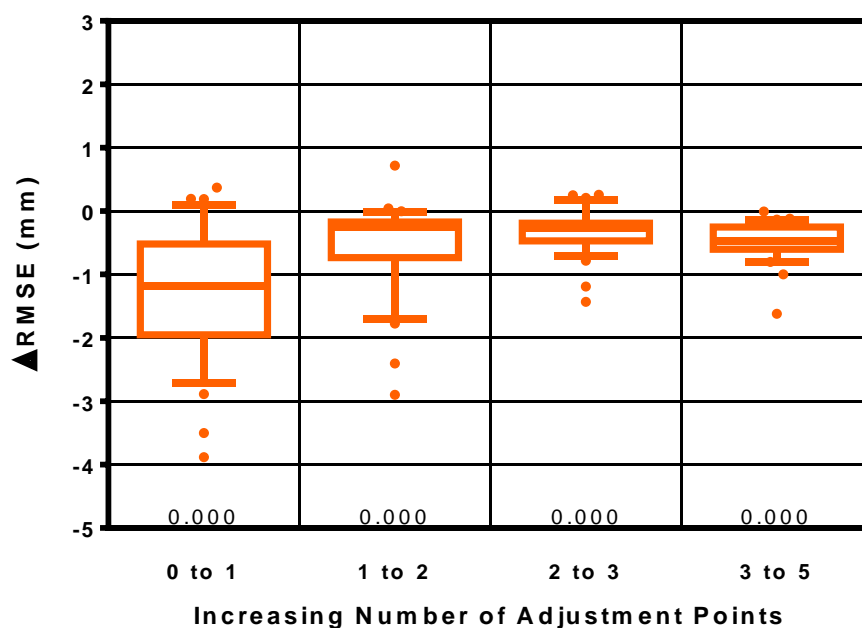


Figure 73: Marginal change in forecast error of the linear model with subsequent adjustment points – Spherical harmonic GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

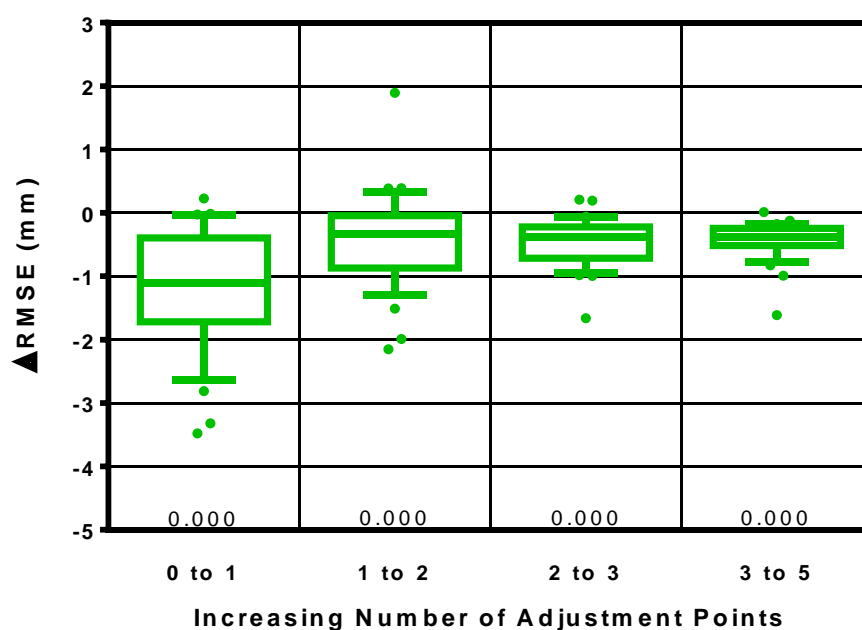


Figure 74: Marginal change in forecast error of the mean model with subsequent adjustment points – Spherical harmonic GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

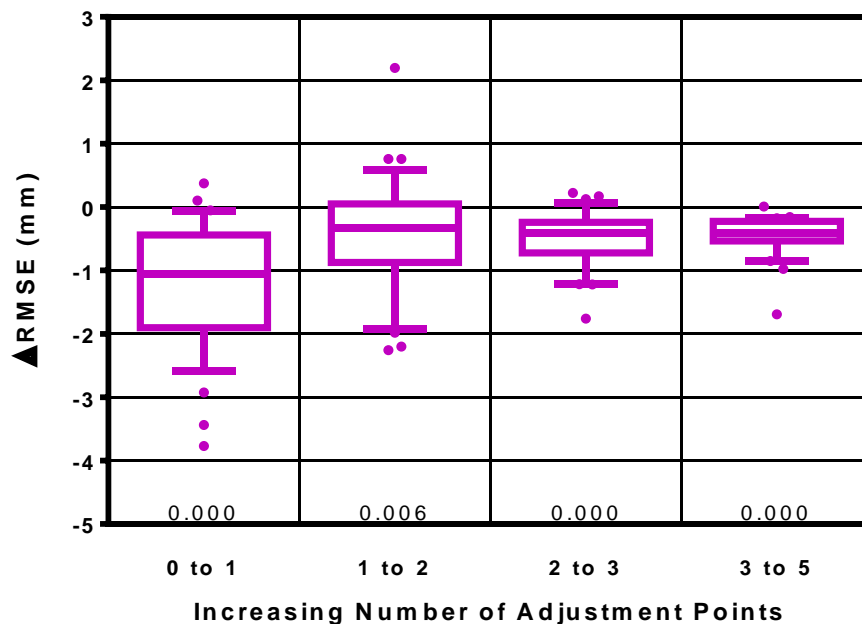


Figure 75: Marginal change in forecast error of the median model with subsequent adjustment points – Spherical harmonic GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

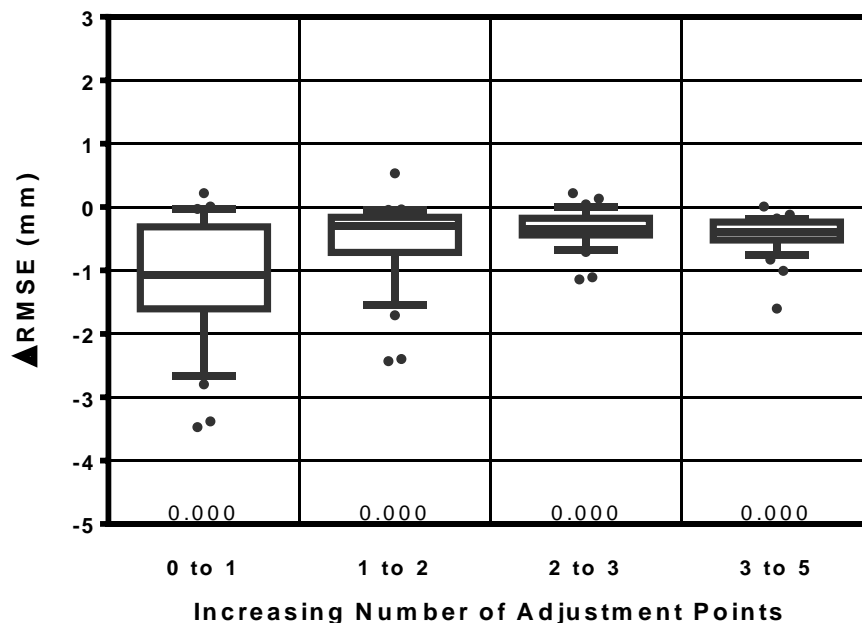


Figure 76: Marginal change in forecast error of the vector PCA model with subsequent adjustment points – Spherical harmonic GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

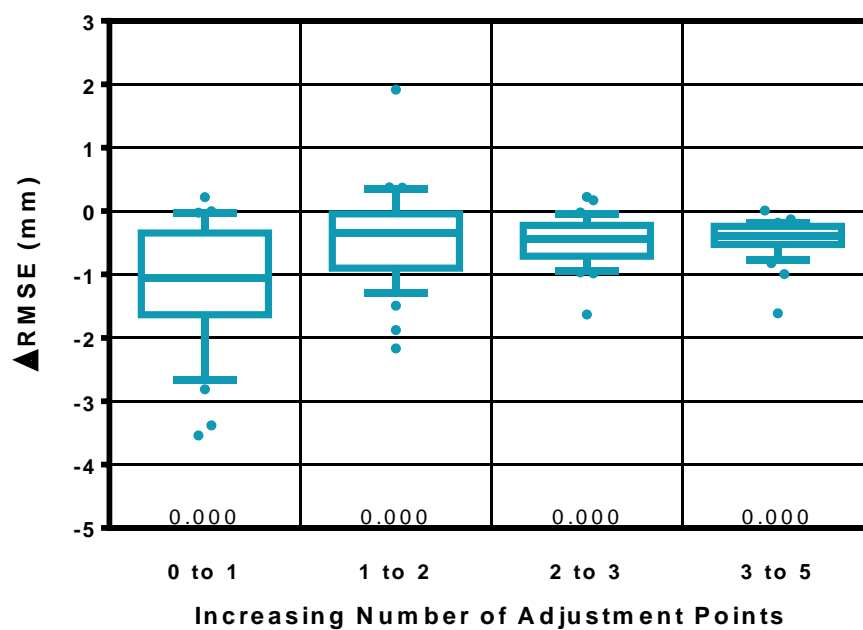


Figure 77: Marginal change in forecast error of the 2D PCA model with subsequent adjustment points – Spherical harmonic GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

4.2.3. Specific Aim 3 results

Specific Aim 3 pertains to the effect of the model type on the forecast error for models of GTV volume and GTV centroid position.

4.2.3.1. Specific Aim 3 results – GTV volume descriptor

Figures 78 – 82 depict the change in forecast error due to selection of the non-static models of GTV volume for each number of adjustment points. With only two exceptions (the median and 2D PCA models with 5 adjustment points), the five non-static models always resulted in a statistically significant decrease in forecast error. The range and magnitude of these changes diminished with the increasing number of adjustment points. For models with no adjustment points, the non-static models decreased the error compared to the static model by 11.7% – 12.5%. These changes corresponding to a 54.2% – 58.0% decrease relative to the median error of the static model.

For models of GTV volume, the relative decrease in forecast error achieved by selection of each non-static model was considerably larger than that achieved with the addition of adjustment points in Specific Aim 1.

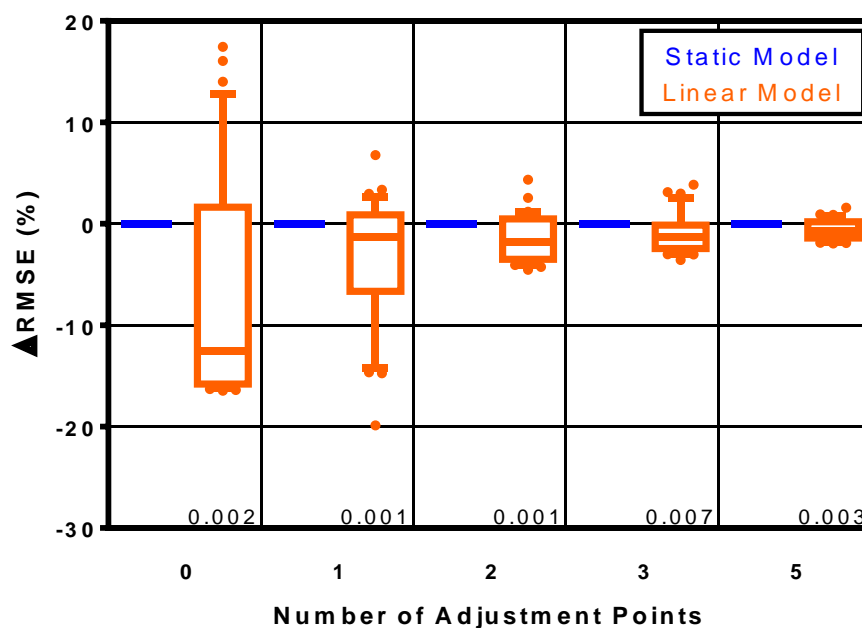


Figure 78: Change in forecast error from the static model to the linear model with an equal number of adjustment points – GTV volume descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

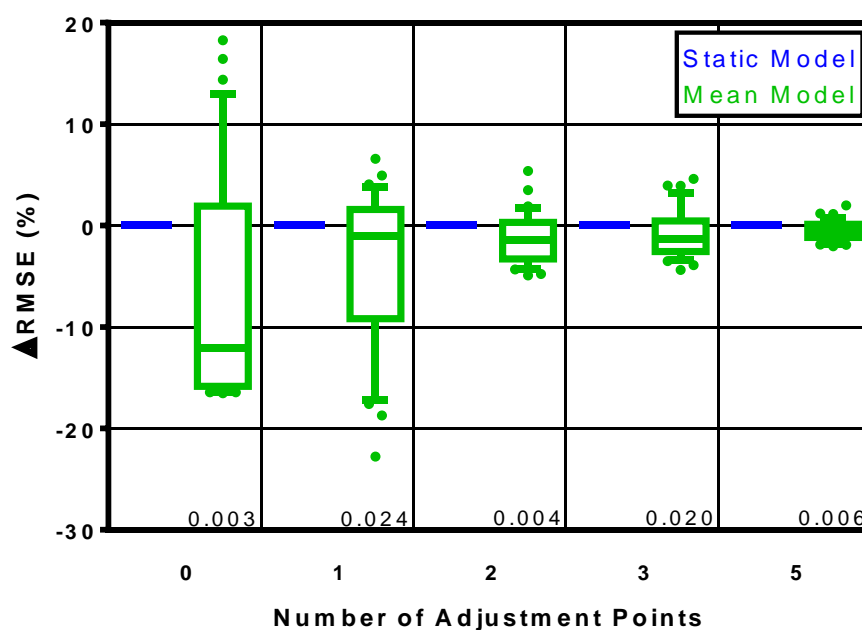


Figure 79: Change in forecast error from the static model to the mean model with an equal number of adjustment points – GTV volume descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

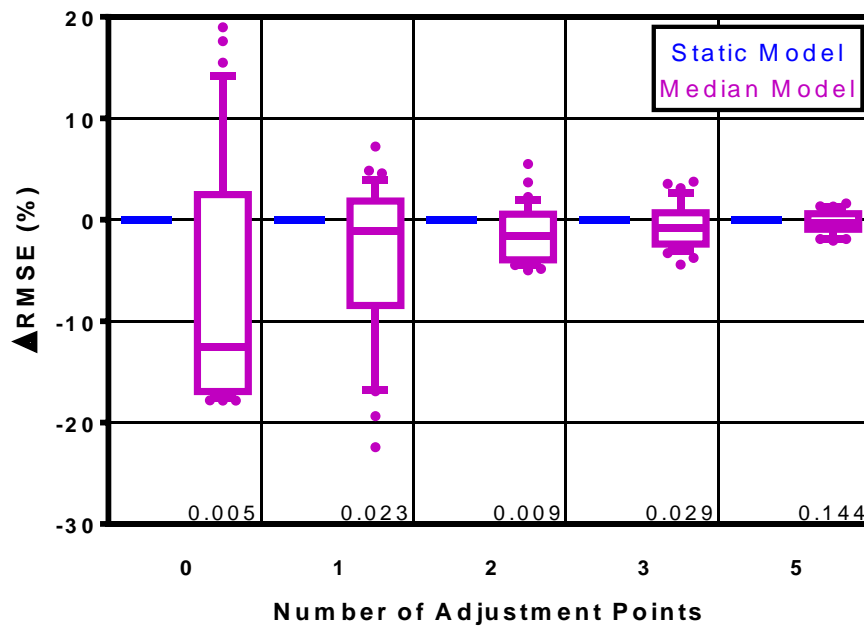


Figure 80: Change in forecast error from the static model to the median model with an equal number of adjustment points – GTV volume descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

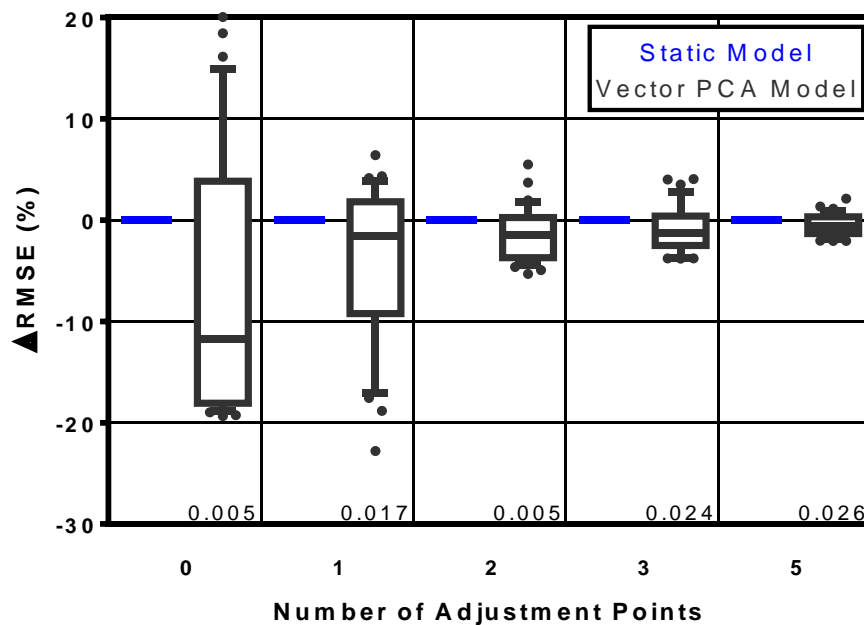


Figure 81: Change in forecast error from the static model to the vector PCA model with an equal number of adjustment points – GTV volume descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

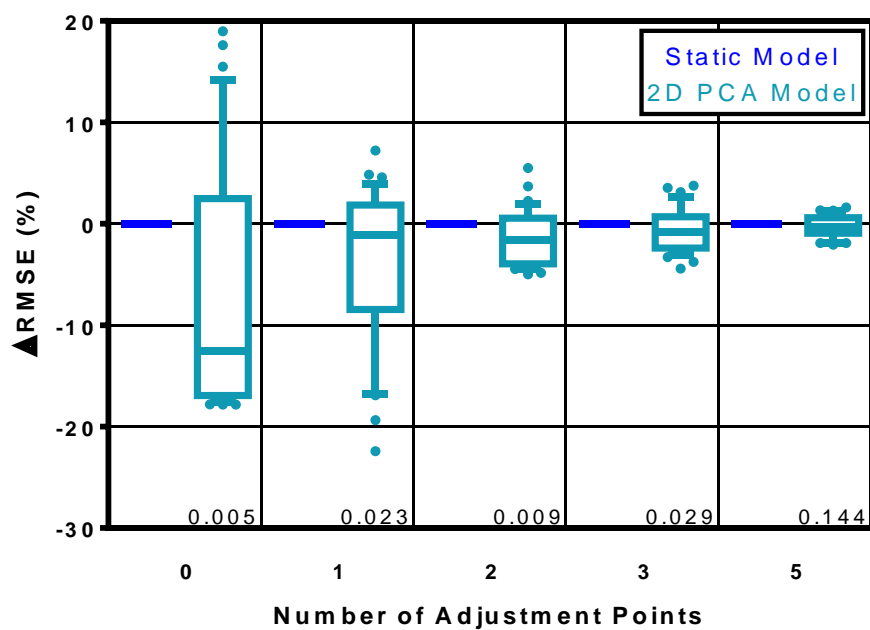


Figure 82: Change in forecast error from the static model to the 2D PCA model with an equal number of adjustment points – GTV volume descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

4.2.3.2. Specific Aim 3 results – GTV centroid position descriptor

The change in forecast error due to the selection of non-static models of GTV centroid position is depicted in Figures 83 – 87. Only a subset of GTV centroid position non-static models resulted in a statistically significant change in forecast error compared to the static model. These included a decrease in error for the mean and vector PCA models with 2 adjustment points, and an increase in error for the linear model with 3 adjustment points, the median models with 3 and 5 adjustment points, and the 2D PCA models with 3 and 5 adjustment points. These changes in forecast error were 0.0 mm – 0.2 mm representing changes of 0.9% – 5.1% relative to the median error of the corresponding static model. However, most changes were not statistically significant.

Unlike the comparison between models of GTV volume, the relative decrease in forecast error achieved by selection of each non-static model was considerably smaller than that achieved with the addition of adjustment points in Specific Aim 1.

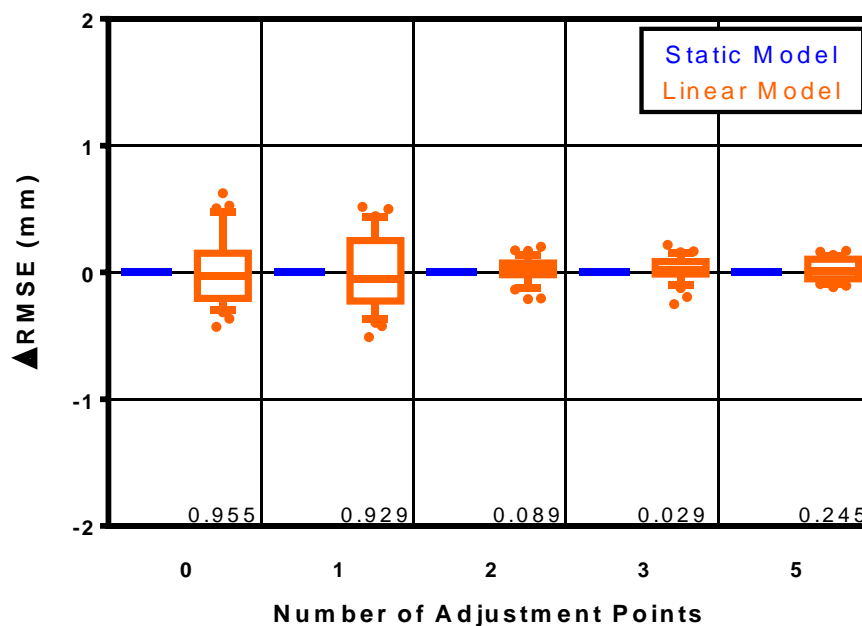


Figure 83: Change in forecast error from the static model to the linear model with an equal number of adjustment points – GTV centroid position descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

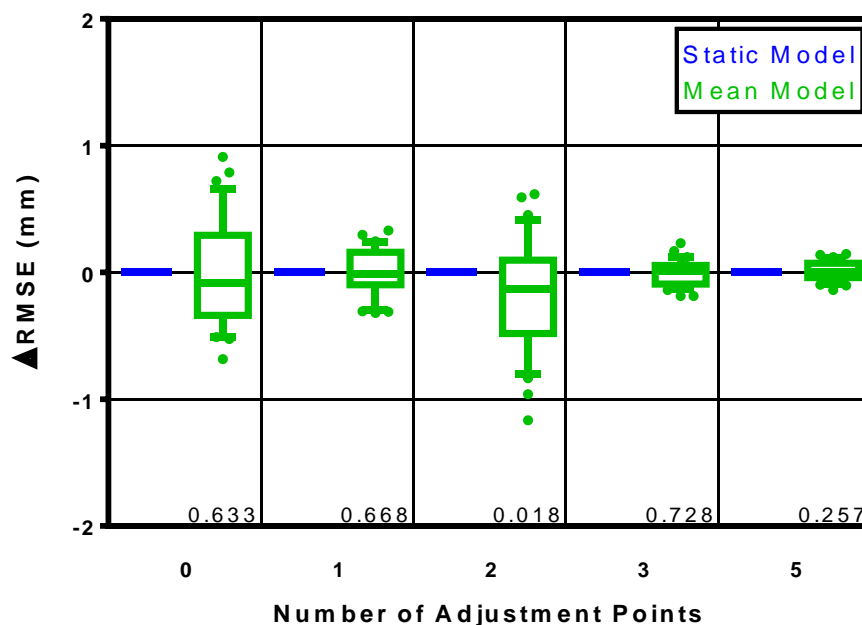


Figure 84: Change in forecast error from the static model to the mean model with an equal number of adjustment points – GTV centroid position descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

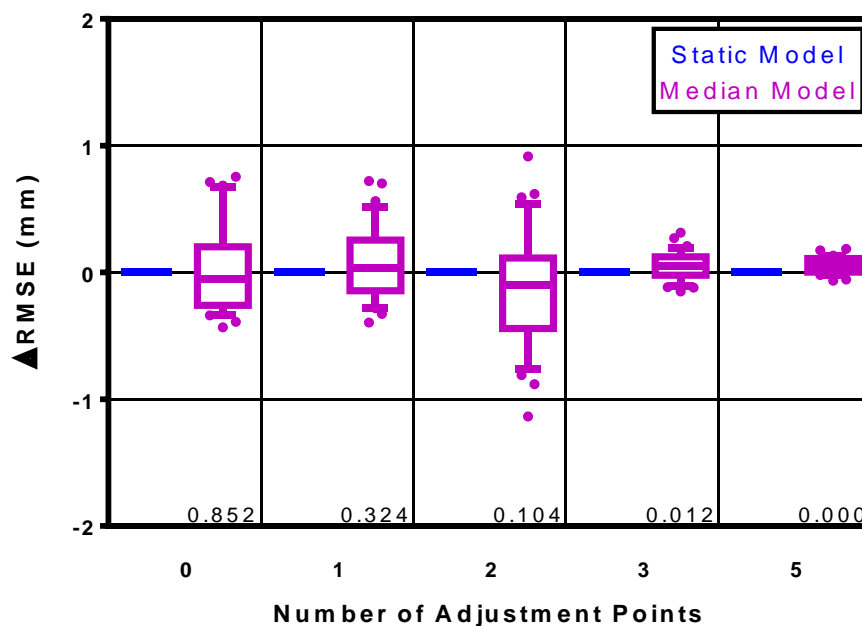


Figure 85: Change in forecast error from the static model to the median model with an equal number of adjustment points – GTV centroid position descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

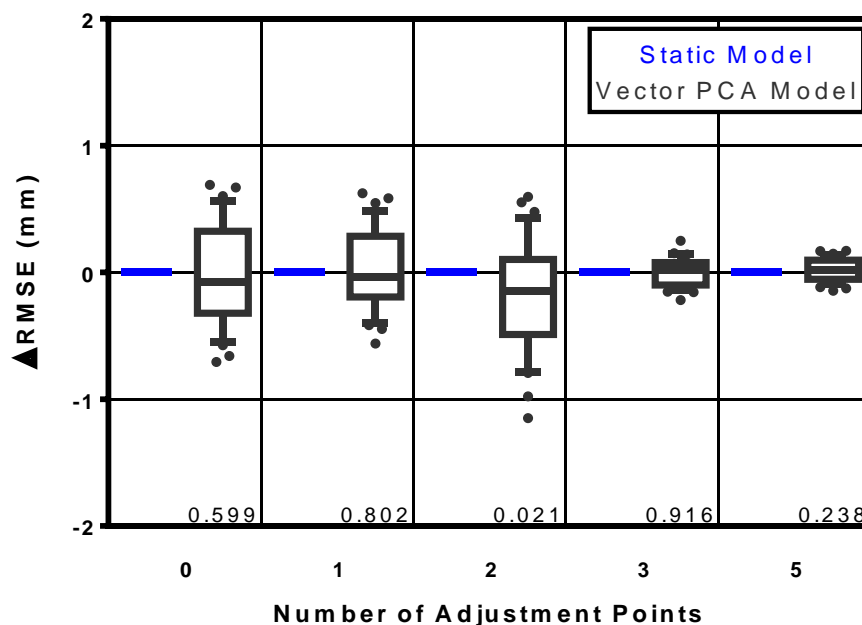


Figure 86: Change in forecast error from the static model to the vector PCA model with an equal number of adjustment points – GTV centroid position descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

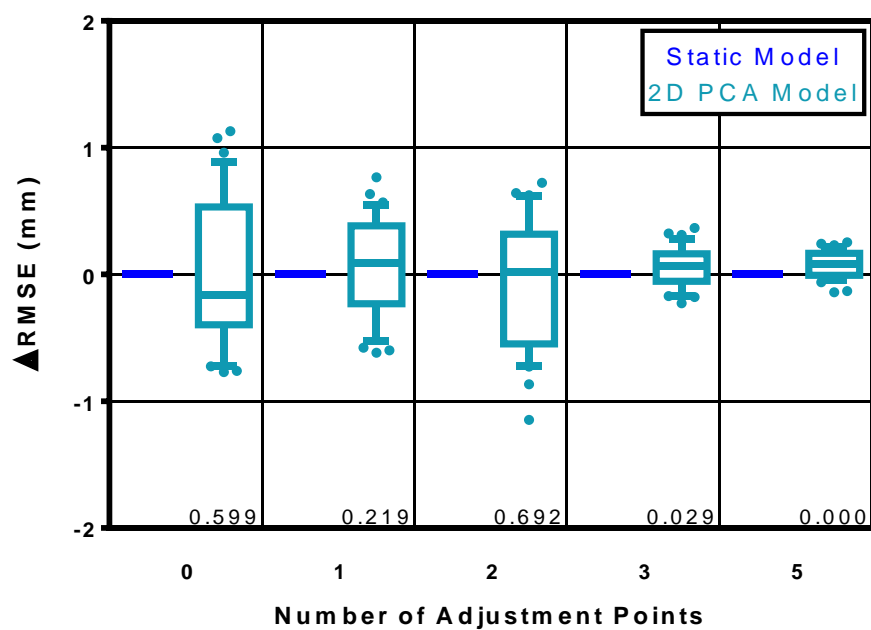


Figure 87: Change in forecast error from the static model to the 2D PCA model with an equal number of adjustment points – GTV centroid position descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

4.2.4. Specific Aim 4 results

The comparisons of Specific Aim 4 are similar to those of Specific Aim 3, but determine the effect of model type on the forecast error for the radial extent and spherical harmonic GTV morphology descriptors.

4.2.4.1. Specific Aim 4 results – Radial extent GTV morphology descriptor

Figures 88 – 92 depict the change in forecast error due to selection of the non-static models of the radial extent GTV morphology descriptor. For models with no adjustment points, selection of each non-static model resulted in a statistically significant decrease in the error compared to the static model. These changes were all 0.2 mm and represented improvements of 3.8% – 5.9% relative to the median error of the corresponding static model.

The relative decrease in forecast error achieved by selection of each non-static model of the radial extent descriptor with no adjustment points was considerably smaller than that achieved with the addition of adjustment points in Specific Aim 2.

Other statistically significant changes were a decrease for the mean model with 1 adjustment point, increases in the median models with 2 and 5 adjustment points, and an increase for the 2D PCA model with 5 adjustment points. These changes were small (0.0 mm – 0.2 mm) representing 1.0% – 5.4% of the median error of the corresponding static model.

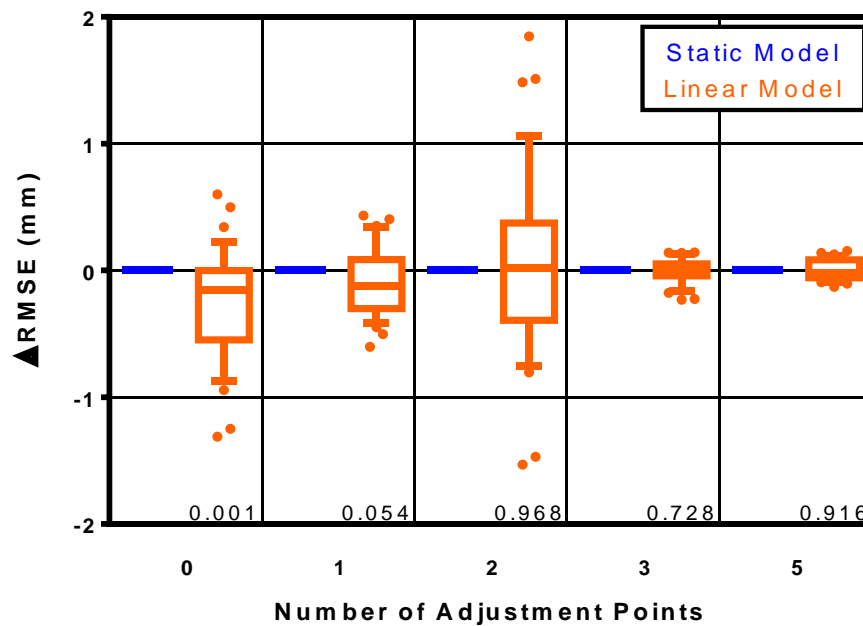


Figure 88: Change in forecast error from the static model to the linear model with an equal number of adjustment points – Radial extent GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

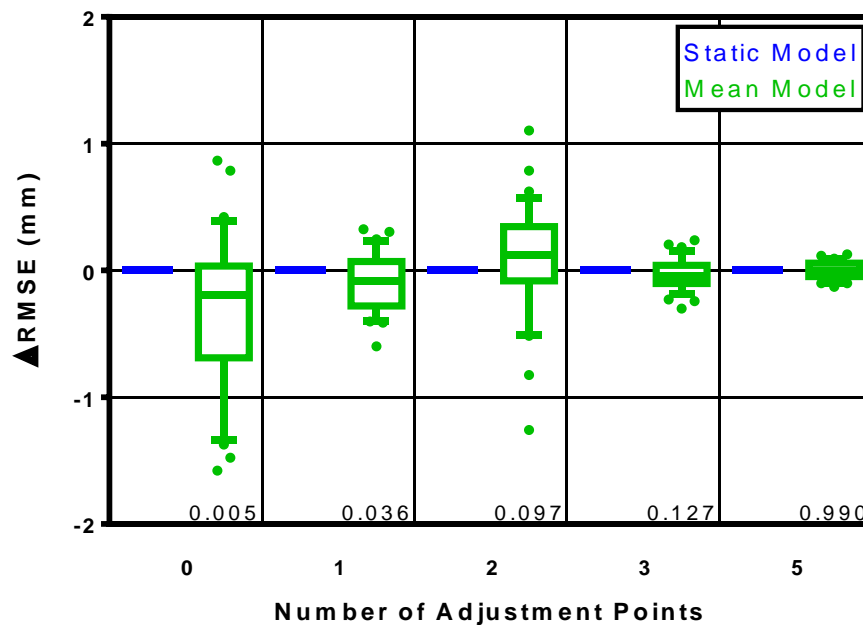


Figure 89: Change in forecast error from the static model to the mean model with an equal number of adjustment points – Radial extent GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

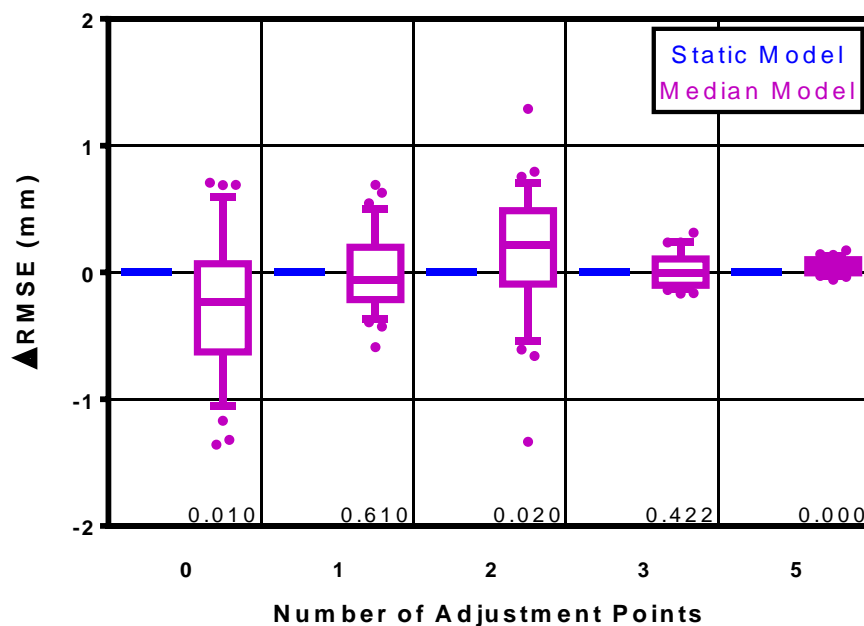


Figure 90: Change in forecast error from the static model to the median model with an equal number of adjustment points – Radial extent GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

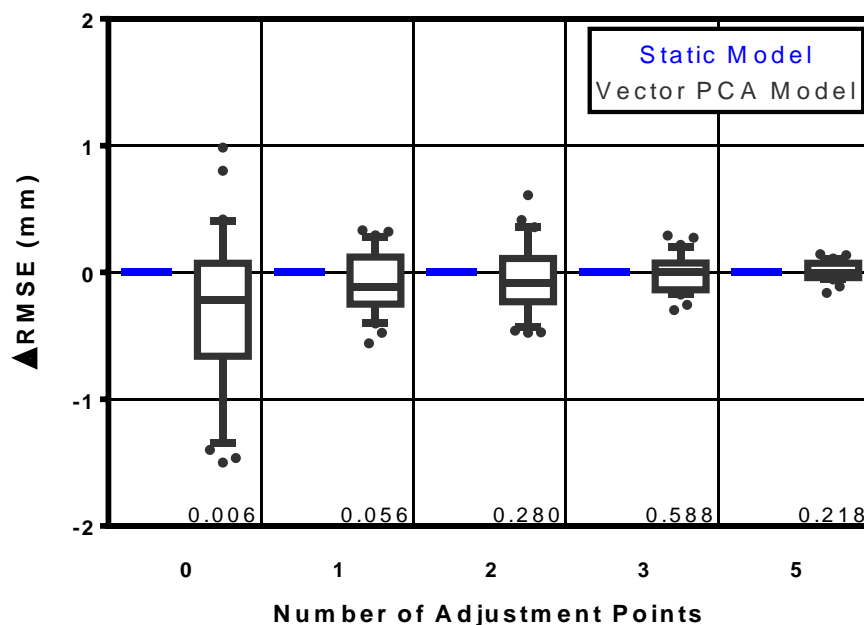


Figure 91: Change in forecast error from the static model to the vector PCA model with an equal number of adjustment points – Radial extent GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

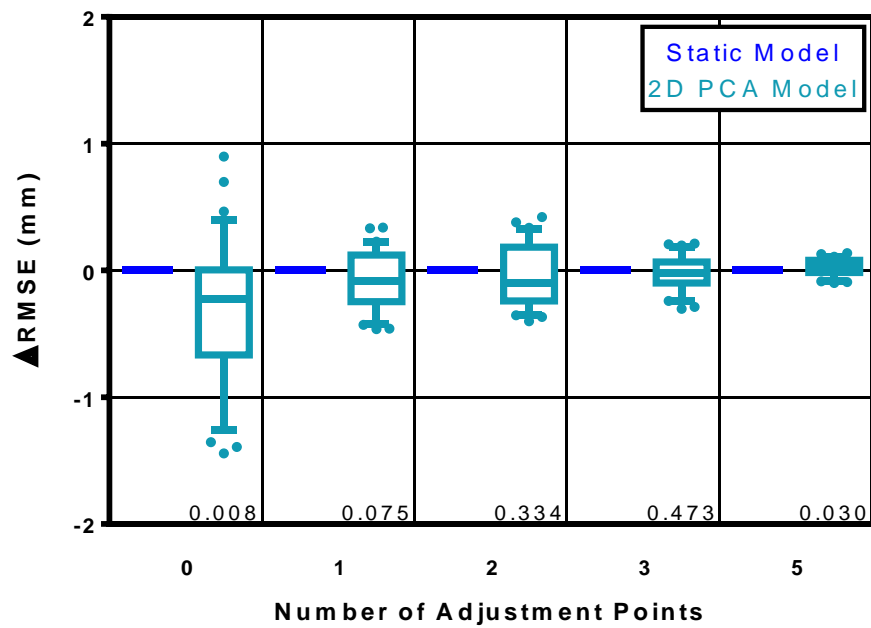


Figure 92: Change in forecast error from the static model to the 2D PCA model with an equal number of adjustment points – Radial extent GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

4.2.4.2. Specific Aim 4 results – Spherical harmonic GTV morphology descriptor

The change in forecast error due to the selection of non-static models of the spherical harmonic GTV morphology descriptor is depicted in Figures 93 – 97. For this descriptor, only the linear and median models with no adjustment points, and the linear model with 1 adjustment point resulted in a statistically significant decrease in error. These changes were 0.2 mm, 0.2 mm and 0.1 mm, respectively, representing changes of 4.7%, 5.1% and 3.6% relative to the median error of the corresponding static model.

Meanwhile, several non-static models featuring a larger number of adjustment points demonstrated a statistically significant increase in the forecast error. These included the mean models with 2, 3 and 5 adjustment points, the median models with 2 and 5 adjustment points, the vector PCA models with 3 and 5 adjustment points, and the 2D PCA models with 2, 3, or 5 adjustment points. These changes were 0.1 mm – 0.3 mm representing 1.2% – 7.3% of the median error of the corresponding static model.

The relative decrease in forecast error achieved by selection of the linear or median models of the spherical harmonic descriptor with no adjustment points was considerably smaller than that achieved with the addition of adjustment points in Specific Aim 2.

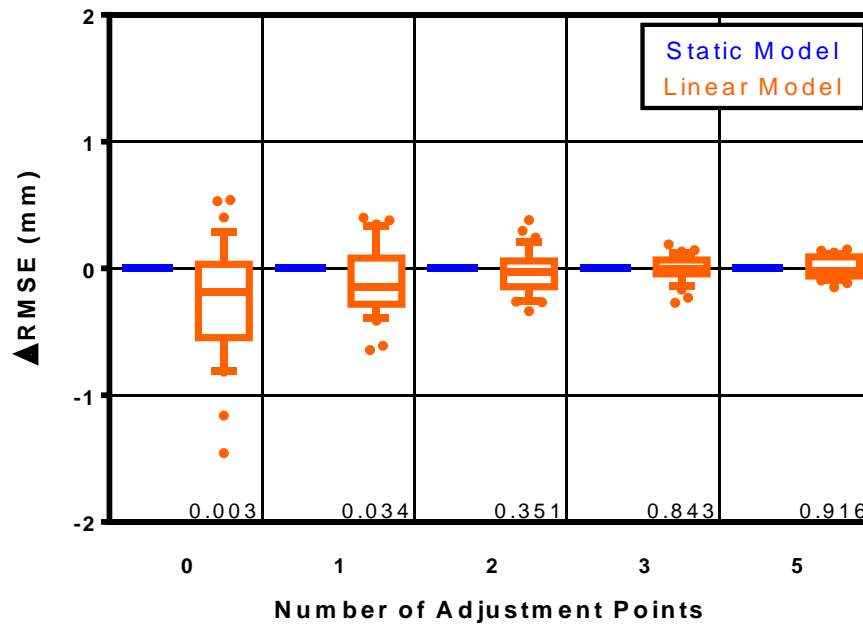


Figure 93: Change in forecast error from the static model to the linear model with an equal number of adjustment points – Spherical harmonic GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

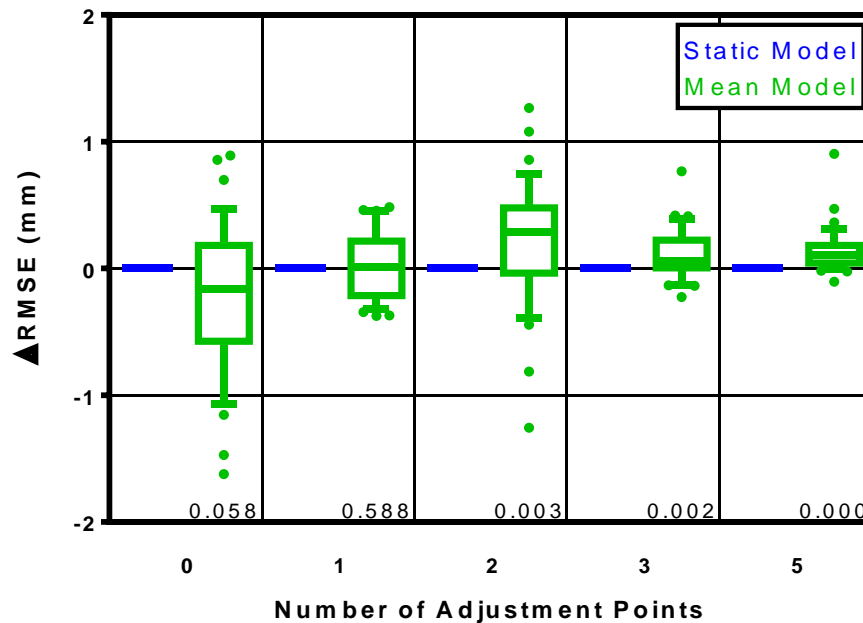


Figure 94: Change in forecast error from the static model to the mean model with an equal number of adjustment points – Spherical harmonic GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

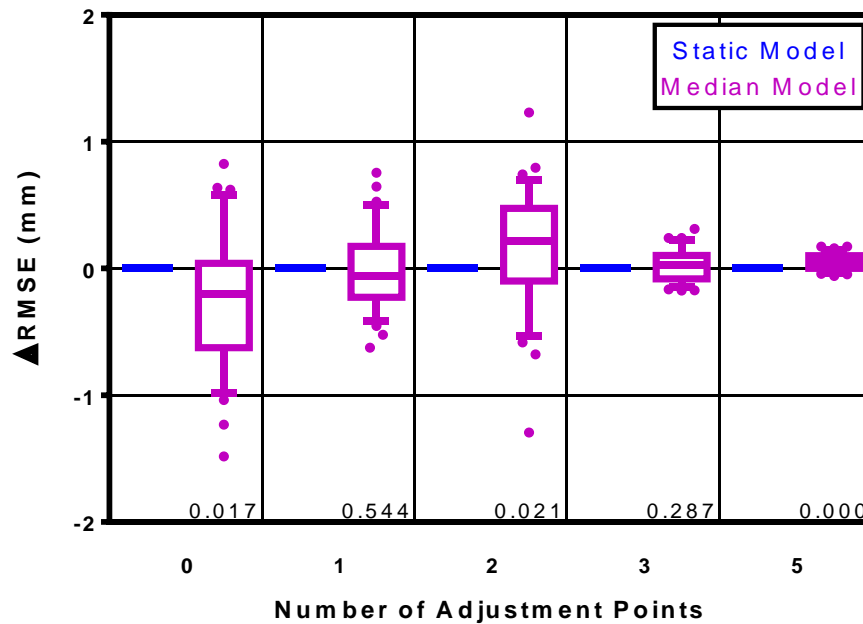


Figure 95: Change in forecast error from the static model to the median model with an equal number of adjustment points – Spherical harmonic GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

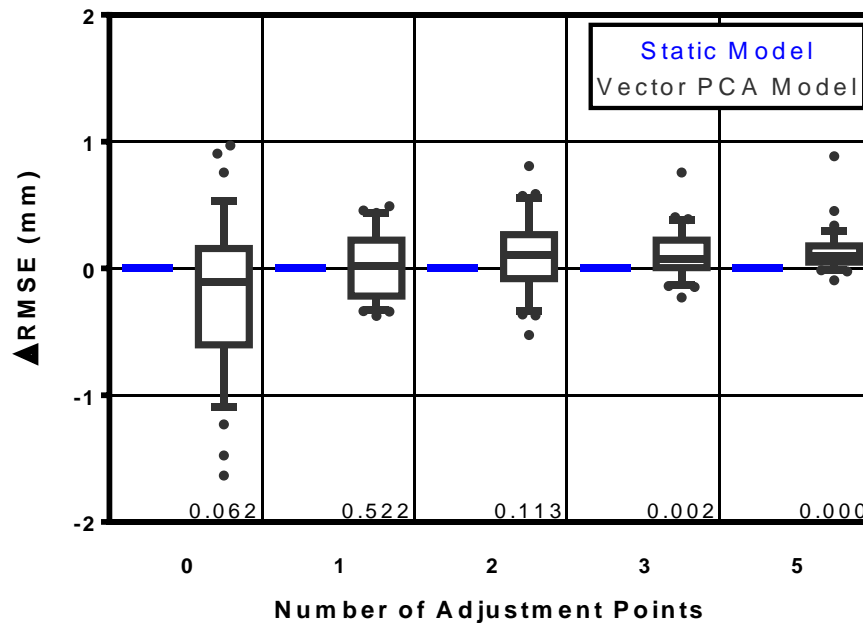


Figure 96: Change in forecast error from the static model to the vector PCA model with an equal number of adjustment points – Spherical harmonic GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

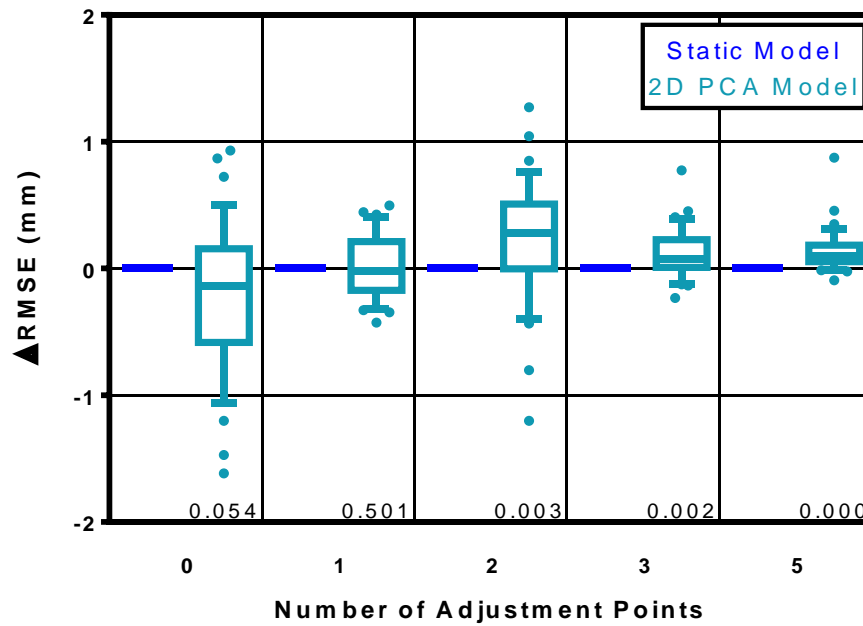


Figure 97: Change in forecast error from the static model to the 2D PCA model with an equal number of adjustment points – Spherical harmonic GTV morphology descriptor. Statistical p-values are included below each box. Box range is 25th to 75th percentile; whisker range is 10th to 90th percentile.

4.3. Reconstructed forecast morphology

To visualize the analysis above, forecasts of the spherical harmonic GTV morphology descriptor from select models were imported into a treatment planning system and compared with the actual anatomy and tumor morphology of that fraction for a single example patient.

In Figure 98, the forecast for the 32nd treatment fraction according to the static models with 0, 1, and 5 adjustment points is depicted with respect to the patient anatomy of that fraction. The blue contour of the forecast from the static model with 5 adjustment points conforms to the actual morphology (represented by the sky blue color wash) more closely than the yellow contour of the static model with 1 adjustment point. The yellow contour, in turn, more closely approximates the actual tumor morphology than the red contour of the static model with no adjustment points. This matches the observation that the forecast accuracy improves with the increasing number of adjustment points. A region of the GTV may not monotonically approach the actual morphology because the model accuracy was measured for the morphology as a whole.

In Figure 99, the forecast for the 32nd treatment fraction according to the static and linear models with no adjustment points is depicted with respect to the patient anatomy of that fraction. The blue contour of the forecast from the linear model conforms to the actual morphology more closely than the red contour of the static model. This improvement is most apparent in Figure 99A where the static model forecast extends substantially into the hyoid bone. Even for the morphology forecast from the linear model, some discrepancies remain. But imperfect conformality is to be expected because, with no adjustment points, neither forecast incorporated patient-specific information acquired during treatment. These predictions were therefore, in effect, determined at the time of the original planning CT.

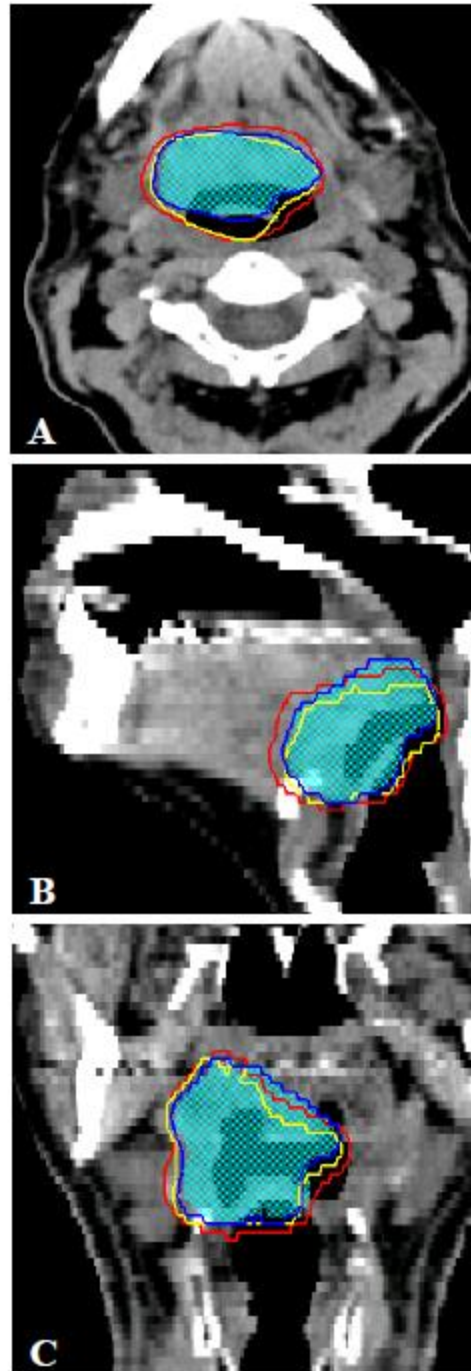


Figure 98: Forecast contours of static models with different numbers of adjustment points. The forecasts are depicted on an (A) axial image, (B) sagittal image, and (C) coronal image of the anatomy observed at treatment fraction 32. Sky blue color wash: original GTV contour deformed to treatment fraction 32, considered ground truth. Red contour: forecast morphology according to the static model with no adjustment points. Yellow contour: forecast morphology according to the static model with a single adjustment point at fraction 16. Dark blue contour: forecast morphology according to the static model with 5 adjustment points.

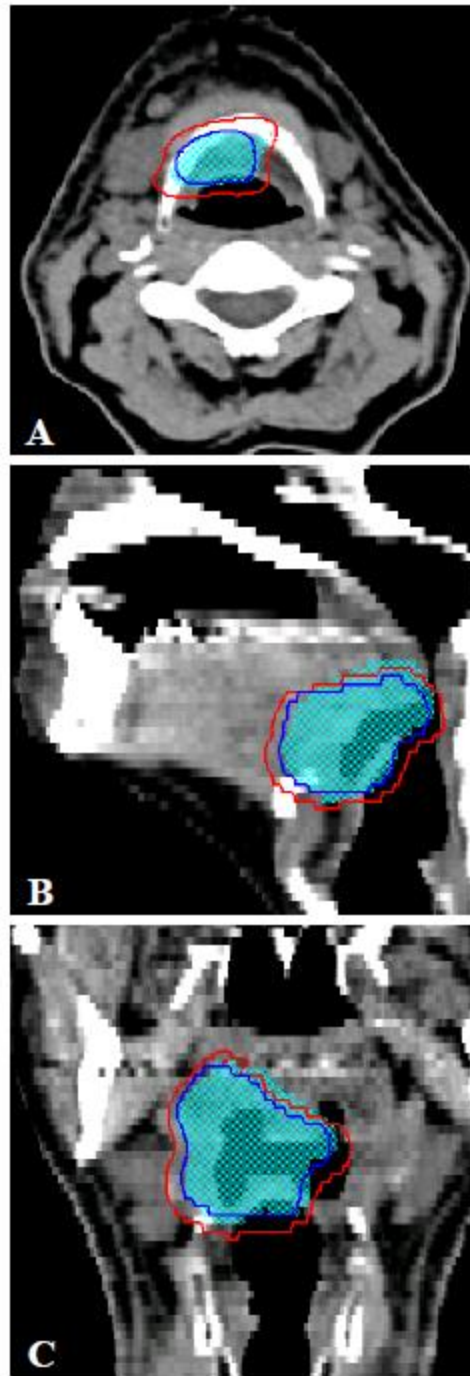


Figure 99: Forecast contours of static and linear models with no adjustment points. The forecasts are depicted on an (A) axial image, (B) sagittal image, and (C) coronal image of the anatomy observed at treatment fraction 32. Sky blue color wash: original GTV contour deformed to treatment fraction 32, considered ground truth. Red contour: forecast morphology according to the static model with no adjustment points. Dark blue contour: forecast morphology according to the linear model with no adjustment points.

Chapter 5

DISCUSSION

The accuracy in forecasting tumor volume, position, and morphology throughout radiation therapy was evaluated using several model types and model parameter adjustment schemes. Non-static models and models that included parameter adjustment points were both observed to improve the ability to predict these tumor variables. These effects are logical. Parameters of the non-static models were determined based on similar patients, incorporating population-based prior information into the predictions. Meanwhile, adjusting model parameters to match the observed response of the patient being treated incorporated patient-specific prior information. Each of the four descriptor variables that were studied exhibited its own dependence of forecast accuracy on model type and number of adjustment points. These relationships are discussed individually below.

The GTV volume descriptor demonstrated an improvement in forecast accuracy for models that included any number of adjustment points. Relative to models with no adjustment points, those with a single adjustment point provided a 12.9% – 30.0% improvement. This exceeded the 10% improvement stated in the hypothesis of Specific Aim 1. The hypothesis was therefore accepted as it pertained to the GTV volume descriptor. Forecast accuracy continued to improve with additional adjustment points. This was also logical as the models were being adjusted to match the correct feature vector values with increasing frequency.

For models of the GTV volume with no adjustment points, non-static models provided a 54.2% – 58.0% improvement compared to the static model. This far

exceeded the 10% improvement stated in the hypothesis for Specific Aim 3. Therefore this hypothesis, too, was accepted with respect to the GTV volume descriptor.

Among non-static model types, there did not appear to be substantial variation in the forecast error due to type. However, the large number of direct statistical comparisons was not performed. With respect to model type selection, the forecast accuracy was improved simply by selecting any one of the non-static models.

The linear model was the simplest non-static model with the fewest parameters and required the least amount of information from the training dataset. Nonetheless, it achieved forecast accuracy comparable to that of the other non-static models. This suggests that the added complexity of the other non-static models was not necessary. That the flexibility of the other non-static models, originally considered capable of modeling more nuanced changes in GTV volume, did not translate into improved forecasts could be due to a number of effects. The more complex models may not have been able to realize their potential of depicting non-linear trends across the training dataset if sufficient variation existed in the changing tumor volumes. In addition, lacking the flexibility of the more complex models may actually have improved the forecast accuracy of the linear model. The constraint of a constant rate of change could have made the linear model more robust to variations among the training dataset, and thus more generalizable when applied to a new patient.

Alternatively, changes in GTV volume may have been modeled by the set of linear models sufficiently accurately so that non-linear models failed to add substantial benefit. Barker et al. considered the changing volume of oropharyngeal tumors as linear, but observed the rate of change was not constant.¹⁸ Non-constant rates of change

could effectively be described using the linear models presented here with the inclusion of adjustment points. The ability to adjust the rate of change at one or more points might provide sufficient accuracy and render the added complexity of the non-linear models unnecessary.

For the GTV volume descriptor, the improvement in forecast accuracy from selection of a non-static model type was greater than that achieved by adding an adjustment point. However, incorporating both strategies had an even greater effect on forecast accuracy.

Considering all model types, the addition of a single adjustment point to models of the GTV centroid position improved the forecast accuracy by 27.0% – 34.1% relative to models with no adjustment points. This was greater than the 10% stated in the hypothesis for Specific Aim 1. Therefore, the hypothesis of Specific Aim 1 was accepted with respect to GTV centroid position descriptor, as it was for the GTV volume descriptor.

Unlike for the GTV volume descriptor, however, the marginal improvement in forecast accuracy due to adding subsequent adjustment points clearly diminished with each successive adjustment point. This effect is again logical as the increased number of adjustment points limited the opportunity for the predicted values to drastically diverge from the actual values. This diminishing effect is important because the clinical implementation of such adjustment points is not trivial. In generating the models presented here, parameters were simply adjusted to match recently observed values of the feature vectors. Clinically, this corresponds to adapting a patient's treatment. This requires the patient to receive additional imaging exams costing time and money as

well as exposure to greater radiation dose. Furthermore, altering the patient's treatment in order to adapt to the more recent images requires significant time and effort on behalf of the clinicians. These costs must be weighed against the perceived potential clinical benefit.

That the improvement in GTV centroid position forecast accuracy diminished with additional adjustment points suggests that there may be an optimal number of adjustment points – one where the improvement in accuracy outweighs the cost of adjusting the treatment, while additional adjustments would not be justified. It is beyond the scope of this work to determine the optimal number of patient treatment adjustments. This is clinical decision that requires the expertise of the radiation oncologist and consideration of the individual patient.

Model type did not have a large impact on forecast accuracy for models of GTV centroid position. Most of the differences between non-static and static models with an equal number of adjustment points were not statistically significant. The differences that were significant (0.9% – 5.1% relative to the corresponding static model) were less than the 10% stated in the hypothesis for Specific Aim 3. Therefore, unlike for the GTV volume descriptor, the hypothesis of Specific Aim 3 was rejected for the GTV centroid position descriptor. Adding adjustment points proved to be the more effective way to improve forecast accuracy for this descriptor.

For all model types, adding adjustment points to models of the radial extent GTV morphology descriptor improved the forecast accuracy by 27.5% – 33.8% compared to models with no adjustment points. This was greater than the 10% improvement stated in the hypothesis for Specific Aim 2, which was therefore accepted. As observed for

models of the GTV centroid position, the marginal improvement in forecast accuracy for models of the radial extent descriptor diminished with each additional adjustment point. This again suggests that the improvement achieved with the addition of some adjustment points may not always be worth the effort of adjusting a patient's treatment.

For models with no adjustment points, non-linear models provided a 3.8% – 5.9% improvement in forecast accuracy compared to the static model. This was less than the 10% required to accept the hypothesis of Specific Aim 4, which was therefore rejected with respect to the radial extent descriptor.

The improvement in forecast accuracy acquired through the addition of adjustment points was considerably greater than that provided by the non-static models. This implies that for the radial extent descriptor, a greater clinical benefit is to be achieved by re-imaging the patient and re-optimizing the treatment plan than by assuming a non-static model type. Though small, there was an additional benefit of using a non-static model type in addition to adjusting the model parameters.

The forecast accuracy of models of the spherical harmonic GTV morphology descriptor varied with the number of adjustment points in a way similar to those of the radial extent descriptor. The addition of any number of adjustment points resulted in a significant improvement in forecast accuracy for all model types. Even a single adjustment point resulted in a 28.6% – 33.0% improvement compared to the models with no adjustment points. This was greater than the 10% stated in the hypothesis for Specific Aim 2. The hypothesis was therefore accepted with respect to the spherical harmonic descriptor in addition to the radial extent descriptor. Furthermore, the marginal improvement of successive adjustment point again exhibited a diminishing

effect, suggesting a limit to the number of adjustments that would be clinically justifiable when using these models.

For models of the spherical harmonic descriptor with no adjustment points, only the linear and median models resulted in a significant improvement in forecast accuracy compared to the static model (4.7% and 5.1%, respectively). These improvements were less than 10% needed to accept the hypothesis of Specific Aim 4. This hypothesis was therefore rejected for the spherical harmonic descriptor, as it was for the radial extent descriptor.

It is unclear why the improvement in forecast accuracy offered by non-static models based on the spherical harmonic descriptor was more sensitive to model selection. Both morphology descriptors were capable of closely approximating tumor morphology. One advantage of the spherical harmonic descriptor was that it efficiently represented an entire surface, not just the position of the surface landmarks. However, when generating the spherical harmonic feature vector, the intermediate step of projecting distances onto the spherical harmonic basis may have made this descriptor less appropriate for comparing changes in morphology. The radial extent feature vector directly reflected the original data of the distances between the GTV centroid and surface landmarks. Compared to a feature vector of spherical harmonic coefficients, transferring information regarding changes in these distances was more intuitive and less subject to exaggerated effects of small changes. As a result, the radial extent descriptor may have been more consistent and more useful when not merely describing morphology, but transferring information regarding changes in morphology.

As depicted in Table III, the hypotheses of Specific Aim 1 and Specific Aim 2 were accepted for both reductive descriptors and both generative descriptors, respectively. The hypothesis of Specific Aim 3 was accepted for the GTV volume descriptor but not for the GTV centroid position descriptor, and that of Specific Aim 4 was rejected for both generative descriptors. Overall, these results support the Principal Hypothesis:

Longitudinal trends in the volume, position, and morphology of head and neck tumors can be derived from previous patients and applied to reduce the error in forecasting these variables during the treatment of a new patient by 10% compared to the original planning CT. (Chapter 2)

Although the improvement in forecast accuracy from a non-static model alone was small for each descriptor variable other than GTV volume, the addition of a single adjustment point, either alone or in conjunction with a non-static model, was more than enough to improve the forecast accuracy by the 10% stated in the hypothesis.

Table III: Summary of the results for the four Specific Aims. (✓) or (✗) denotes that the improvement was greater than or less than 10% resulting in the acceptance or rejection of the specific aim hypothesis, respectively.

Effect of:	Reductive Descriptors	Generative Descriptors
Single Adjustment point	<u>Specific Aim 1:</u> ✓ GTV volume ✓ GTV centroid position	<u>Specific Aim 2:</u> ✓ Radial extent GTV morphology ✓ Spherical harmonic GTV morphology
Non-static Model	<u>Specific Aim 3:</u> ✓ GTV volume ✗ GTV centroid position	<u>Specific Aim 4:</u> ✗ Radial extent GTV morphology ✗ Spherical harmonic GTV morphology

Tumors analyzed in this work were derived from patients on a clinical protocol. The patients were required to meet the protocol's selection criteria, and therefore had very similar diseases and treatments. This homogeneity was important in identifying patterns in treatment response and generating meaningful models. However, the patterns cannot be presumed to apply to dissimilar patients. Patients with different diseases or treatments may respond quite different than those used to derive the models presented here. It is therefore inappropriate to generalize the results presented here and apply them to a different patient population. It is also possible that the treatment response exhibited by these patients is better described by separating the patients into two or more sub-categories. Were these sub-categories identified, and were a patient to be accurately classified, models specific to each sub-category might further improve the forecast accuracy.

Models presented here were based only on tumor volume, position, or morphology. However, many factors contribute to determine how an individual patient responds to treatment, and the inclusion of clinical and dosimetric parameters might further improve forecast accuracy. For example, HPV status has been shown to be a strong, independent prognostic indicator of survival for oropharyngeal cancer patients.⁷⁷ Models that incorporate HPV status might therefore reduce the uncertainty associated with the predicted values of tumor variables.

Of the tumor variables considered, the GTV volume and centroid position are restricted in their application due to their inherent reductive nature. Their utility is primarily to serve as an indicator of significant anatomic change. This may result in re-imaging the patient and altering their treatment management.

The limited application of the reductive descriptor variables was the impetus for considering generative models of tumor morphology. The radial extent and spherical harmonic GTV morphology descriptors were both based on discrete sampling of the GTV surfaces. The 614 surface landmarks only approximated the GTV surface. Nonetheless this approximation was considered sufficient to observe the primary changes in tumor morphology.

Generative descriptors have been used previously to describe tumor shape. However, these applications do not include the prediction of longitudinal changes. Takao et al. measured changes in cervical lymph nodes using finite element models. However, they only allude to predicting these changes indirectly through models of lymph node volumes that assume uniform shrinkage.⁵³ With no predictive models of tumor morphology described previously, the accuracy of the models presented here cannot be compared directly with any other. Models of reductive descriptors and models quantifying random variations in tumor shape are not equivalent, and comparisons with them are not appropriate.

Methods described here are not limited to representing longitudinal changes in tumors of the oropharynx. Radial extent and spherical harmonic descriptors can be used to quantify random or systematic variations in addition to longitudinal ones, though predictive models would no longer be pertinent. In addition to tumors of the oropharynx, these descriptors can also model the morphology of nearby organs at risk, such as the parotid glands, as well as organs of other anatomic sites. Different organs at different sites will exhibit unique anatomic variations. However, as the morphology descriptors simultaneously represent size, shape, and position, they could model the

periodic changes of a lung tumor due to respiratory motion as well as the random changes of a prostate due to filling of the bladder and rectum. Furthermore, as these descriptors fundamentally represent changes in the configuration of a set of landmarks, they need not represent a specific or single organ at all. The relative position of any set of anatomic points of interest could be represented using methods similar to those presented here.

Predictions made by the models presented here are intended to augment, not supplant, images of the patient acquired during the course of treatment. In fact, the results demonstrated the value of considering updated patient-specific information. However, the value of these models is in expanding the understanding of changes in the tumor when such imaging is not available or during the period between such images.

While uncertainty remains regarding predictions of a system as complex as a patient's response to radiation therapy, the ability of these models to describe changes in tumor variables could be used in numerous ways. Prior to the start of treatment for a new patient, predictive models could anticipate the patient's response to treatment. The predictions of the morphology descriptors could be visualized relative to the patient's anatomy or to a proposed dose distribution, which could inform clinicians' decisions regarding patient management. For example, were models to predict tumor motion towards a radiosensitive organ at risk, or large variability in tumor morphology in a region of steep dose gradients, the original treatment plan might be adjusted. The models might also reveal the importance of subsequent imaging along with the appropriate number and timing of related interventions. In principle, it could even be possible to generate treatment plans based on these predictions. This could be a single

plan optimized to the changing tumor morphology, or a set of re-plans to be administered later during treatment. Regardless of how they are precisely used, these plans are generated based on yet un-realized configurations of the tumor and anatomy. This treatment management strategy can therefore be considered a form of anticipatory adaptive radiation therapy.

Applications of these models persist during treatment. For example, the observed changes in anatomy could be compared against the predictions made by the models. The updated patient-specific information could be used to adjust the model forecast, as demonstrated here, or to determine if a patient is responding as expected or desired. If not, clinicians would have the opportunity to intervene.

Lastly, models of these tumor variables remain useful even after treatment is completed. The models can be used to interpolate between any mid-treatment images that were acquired, creating a record that logs the patient's approximate treatment response. As a more complete depiction of patient anatomy and treatment dosimetry, this characterization of treatment response could be compared with those of other patients or combined with other clinical parameters for subsequent research studies.

*Chapter 6***CONCLUSION**

The work presented here demonstrates the ability to predict changes in tumor volume, position, and morphology throughout treatment from a patient's planning CT. Several model types were combined with numerous model adjustment schemes to improve the forecast accuracy compared to conventional descriptions that assume no change. The predictive models can be incorporated into the treatment planning process or used by clinicians to make more informed treatment management decisions. A more complete characterization of treatment response offered by these models can also be incorporated into future research studies. Of course, uncertainty and error remain in predictions of these models. However, the work presented here contributes to the complexity with which the dynamic response of a patient to radiation therapy can be represented. Further investigation to discern the full potential of integrating comprehensive, dynamic measures of anatomic variability into radiation therapy is encouraged.

BIBLIOGRAPHY

- ¹Howlader N, Noone AM, Krapcho M, Garshell J, Neyman N, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Cho H, Mariotto A, Lewis DR, Chen HS, Feuer SJ, and Cronin KA (eds). SEER Cancer Statistics Review, 1975-2010, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2010/, based on November 2012 SEER data submission, posted to the SEER web site, April 2013.
- ²Skarsgard DP, Groome PA, Mackillop WJ, Zhou S, Rothwell D, Dixon PF, O'Sullivan B, Hall SF, and Holowaty EJ, "Cancers of the upper aerodigestive tract in Ontario, Canada, and the United States," *Cancer* **88**, 1728-1738 (2000).
- ³Blot WJ, McLaughlin JK, Winn DM, Austin DF, Greenberg RS, Prestonmartin S, Bernstein L, Schoenberg JB, Stemhagen A, and Fraumeni JF, "Smoking and drinking in relation to oral and pharyngeal cancer," *Cancer Research* **48**, 3282-3287 (1988).
- ⁴Tuyns AJ, Esteve J, Raymond L, Berrino F, Benhamou E, Blanchet F, Boffetta P, Crosignani P, Delmoral A, Lehmann W, Merletti F, Pequignot G, Riboli E, Sanchogarnier H, Terracini B, Zubiri A, and Zubiri L, "Cancer of the larynx hypopharynx, tobacco and alcohol - IARC international case-control study in Turin and Varese (Italy), Zaragoza and Navarra (Spain), Geneva (Switzerland) and Calvados (France)," *International Journal of Cancer* **41**, 483-491 (1988).
- ⁵Hashibe M, Brennan P, Chuang SC, Boccia S, Castellsague X, Chen C, Curado MP, Dal Maso L, Daudt AW, Fabianova E, Fernandez L, Wunsch-Filho V, Franceschi S, Hayes RB, Herrero R, Kelsey K, Koifman S, La Vecchia C, Lazarus P, Levi F, Lence JJ, Mates D, Matos E, Menezes A, McClean MD, Muscat J, Eluf-Neto J, Olshan AF, Purdue M, Rudnai P, Schwartz SM, Smith E, Sturgis EM, Szeszenia-Dabrowska N,

- Talamini R, Wei QY, Winn DM, Shangina O, Pilarska A, Zhang ZF, Ferro G, Berthiller J, and Boffetta P, "Interaction between tobacco and alcohol use and the risk of head and neck cancer: Pooled analysis in the International Head and Neck Cancer Epidemiology Consortium," *Cancer Epidemiology Biomarkers & Prevention* **18**, 541-550 (2009).
- ⁶Lenhard RW, Osteen RT, and Gansler T. *Clinical oncology*. American Cancer Society, Inc., Atlanta, GA 2001.
- ⁷Kreimer AR, Clifford GM, Boyle P, and Franceschi S, "Human papillomavirus types in head and neck squamous cell carcinomas worldwide: A systematic review," *Cancer Epidemiology Biomarkers & Prevention* **14**, 467-475 (2005).
- ⁸Wendt TG, Grabenbauer GG, Rodel CM, Thiel HJ, Aydin H, Rohloff R, Wustrow TPU, Iro H, Popella C, and Schalhorn A, "Simultaneous radiochemotherapy versus radiotherapy alone in advanced head and neck cancer: A randomized multicenter study," *Journal of Clinical Oncology* **16**, 1318-1324 (1998).
- ⁹Calais G, Alfonsi M, Bardet E, Sire C, Germain T, Bergerot P, Rhein B, Tortochaux J, Oudinot P, and Bertrand P, "Randomized trial of radiation therapy versus concomitant chemotherapy and radiation therapy for advanced-stage oropharynx carcinoma," *Journal of the National Cancer Institute* **91**, 2081-2086 (1999).
- ¹⁰Adelstein DJ, Li Y, Adams GL, Wagner H, Kish JA, Ensley JF, Schuller DE, and Forastiere AA, "Intergroup phase III comparison of standard radiation therapy and two schedules of concurrent chemoradiotherapy in patients with unresectable squamous cell head and neck cancer," *Journal of Clinical Oncology* **21**, 92-98 (2003).

- ¹¹Olmi P, Crispino S, Fallai C, Torri V, Rossi F, Bolner A, Amichetti M, Signor M, Taino R, Squadrelli M, Colombo A, Ardizzioia A, Ponticelli P, Franchin G, Minatel E, Gobitti C, Atzeni G, Gava A, Flann M, and Marsoni S, "Locoregionally advanced carcinoma of the oropharynx: Conventional radiotherapy vs. accelerated hyperfractionated radiotherapy vs. concomitant radiotherapy and chemotherapy - A multicenter randomized trial," *International Journal of Radiation Oncology Biology Physics* **55**, 78-92 (2003).
- ¹²Denis F, Garaud P, Bardet E, Alfonsi M, Sire C, Germain T, Bergerot P, Rhein B, Tortochaux J, and Calais G, "Final results of the 94-01 French Head and Neck Oncology and Radiotherapy Group randomized trial comparing radiotherapy alone with concomitant radiochemotherapy in advanced-stage oropharynx carcinoma," *Journal of Clinical Oncology* **22**, 69-76 (2004).
- ¹³Huguenin P, Beer KT, Allal A, Rufibach K, Friedli C, Davis JB, Pestalozzi B, Schmid S, Thoni A, Ozsahin M, Bernier J, Topfer M, Kann R, Meier UR, Thum P, Bieri S, Notter M, Lombriser N, and Glanzmann C, "Concomitant cisplatin significantly improves locoregional control in advanced head and neck cancers treated with hyperfractionated radiotherapy," *Journal of Clinical Oncology* **22**, 4665-4673 (2004).
- ¹⁴Budach V, Stuschke M, Budach W, Baumann M, Geismar D, Grabenbauer G, Lammert I, Jahnke K, Stueben G, Herrmann T, Bamberg M, Wust P, Hinkelbein W, and Wernecke KD, "Hyperfractionated accelerated chemoradiation with concurrent fluorouracil-mitomycin is more effective than dose-escalated hyperfractionated accelerated radiation therapy alone in locally advanced head and neck cancer: Final results of the Radiotherapy Cooperative Clinical Trials Group of the German Cancer

- Society 95-06 prospective randomized trial," *Journal of Clinical Oncology* **23**, 1125-1135 (2005).
- ¹⁵Semrau R, Mueller RP, Stuetzer H, Staar S, Schroeder U, Guntinas-Lichius O, Kocher M, Eich HT, Dietz A, Flentje M, Rudat V, Volling P, Schroeder M, and Eckel HE, "Efficacy of intensified hyperfractionated and accelerated radiotherapy and concurrent chemotherapy with carboplatin and 5-fluorouracil: Updated results of a randomized multicentric trial in advanced head-and-neck cancer," *International Journal of Radiation Oncology Biology Physics* **64**, 1308-1316 (2006).
- ¹⁶Leong J, "Implementation of random positioning error in computerized radiation treatment planning systems as a result of fractionation," *Physics in Medicine and Biology* **32**, 327-334 (1987).
- ¹⁷van Herk M, "Errors and margins in radiotherapy," *Seminars in Radiation Oncology* **14**, 52-64 (2004).
- ¹⁸Barker JL, Garden AS, Ang KK, O'Daniel JC, Wang H, Court LE, Morrison WH, Rosenthal DI, Chao KSC, Tucker SL, Mohan R, and Dong L, "Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system," *International Journal of Radiation Oncology Biology Physics* **59**, 960-970 (2004).
- ¹⁹de Boer HCJ and Heijmen BJM, "A protocol for the reduction of systematic patient setup errors with minimal portal imaging workload," *International Journal of Radiation Oncology Biology Physics* **50**, 1350-1365 (2001).
- ²⁰Lee N, Xia P, Fischbein NJ, Akazawa P, Akazawa C, and Quivey JM, "Intensity-modulated radiation therapy for head-and-neck cancer: The UCSF experience focusing on target

- volume delineation," *International Journal of Radiation Oncology Biology Physics* **57**, 49-60 (2003).
- ²¹van Lin ENJT, van der Vight L, Huizenga H, Kaanders JHAM, and Visser AG, "Set-up improvement in head and neck radiotherapy using a 3D off-line EPID-based correction protocol and a customised head and neck support," *Radiotherapy and Oncology* **68**, 137-148 (2003).
- ²²Kim S, Akpati HC, Kielbasa JE, Li JG, Liu CR, Amdur RJ, and Palta JR, "Evaluation of intrafraction patient movement for CNS and head & neck IMRT," *Medical Physics* **31**, 500-506 (2004).
- ²³Drabik DM, MacKenzie MA, and Fallone GB, "Quantifying appropriate PTV setup margins: Analysis of patient setup fidelity and intrafraction motion using post-treatment megavoltage computed tomography scans," *International Journal of Radiation Oncology Biology Physics* **68**, 1222-1228 (2007).
- ²⁴O'Daniel JC, Garden AS, Schwartz DL, Wang H, Ang KK, Ahamad A, Rosenthal DI, Morrison WH, Asper JA, Zhang L, Tung SM, Mohan R, and Dong L, "Parotid gland dose in intensity-modulated radiotherapy for head and neck cancer: Is what you plan what you get?" *International Journal of Radiation Oncology Biology Physics* **69**, 1290-1296 (2007).
- ²⁵Polat B, Wilbert J, Baier K, Flentje M, and Guckenberger M, "Nonrigid patient setup errors in the head-and-neck region," *Strahlentherapie Und Onkologie* **183**, 506-511 (2007).
- ²⁶van Kranen S, van Beek S, Rasch C, van Herk M, and Sonke JJ, "Setup uncertainties of anatomical sub-regions in head-and-neck cancer patients after offline CBCT

- guidance," *International Journal of Radiation Oncology Biology Physics* **73**, 1566-1573 (2009).
- ²⁷Wang J, Bai S, Chen NY, Xu F, Jiang XQ, Li Y, Xu QF, Shen YL, Zhang H, Gong YL, Zhong RM, and Jiang QF, "The clinical feasibility and effect of online cone beam computer tomography-guided intensity-modulated radiotherapy for nasopharyngeal cancer," *Radiotherapy and Oncology* **90**, 221-227 (2009).
- ²⁸Den RB, Doemer A, Kubicek G, Bednarz G, Galvin JM, Keane WM, Xiao Y, and Machtay M, "Daily image guidance with cone-beam computed tomography for head-and-neck cancer intensity-modulated radiotherapy: A prospective study," *International Journal of Radiation Oncology Biology Physics* **76**, 1353-1359 (2010).
- ²⁹Castadot P, Lee JA, Geets X, and Gregoire V, "Adaptive radiotherapy of head and neck cancer," *Seminars in Radiation Oncology* **20**, 84-93 (2010).
- ³⁰Yan D, Vicini F, Wong J, and Martinez A, "Adaptive radiation therapy," *Physics in Medicine and Biology* **42**, 123-132 (1997).
- ³¹Mohan R, Zhang XD, Wang H, Kang YX, Wang XC, Liu H, Ang K, Kuban D, and Dong L, "Use of deformed intensity distributions for on-line modification of image-guided IMRT to account for interfractional anatomic changes," *International Journal of Radiation Oncology Biology Physics* **61**, 1258-1266 (2005).
- ³²Hansen EK, Bucci MK, Quivey JM, Weinberg V, and Xia P, "Repeat CT imaging and replanning during the course of IMRT for head-and-neck cancer," *International Journal of Radiation Oncology Biology Physics* **64**, 355-362 (2006).
- ³³Kuo YC, Wu TH, Chung TS, Huang KW, Chao KSC, Su WC, and Chiou JF, "Effect of regression of enlarged neck lymph nodes on radiation doses received by parotid

- glands during intensity-modulated radiotherapy for head and neck cancer," *American Journal of Clinical Oncology-Cancer Clinical Trials* **29**, 600-605 (2006).
- ³⁴Geets X, Tomsej M, Lee JA, Duprez T, Coche E, Cosnard G, Lonneux M, and Gregoire V, "Adaptive biological image-guided IMRT with anatomic and functional imaging in pharyngo-laryngeal tumors: Impact on target volume delineation and dose distribution using helical tomotherapy," *Radiotherapy and Oncology* **85**, 105-115 (2007).
- ³⁵Schwartz DL and Dong L, "Adaptive radiation therapy for head and neck cancer – Can an old goal evolve into a new standard?" *J Oncol* **2011** (2011).
- ³⁶Schwartz DL, Garden AS, Thomas J, Chen YP, Zhang YB, Lewin J, Chambers MS, and Dong L, "Adaptive radiotherapy for head-and-neck cancer: Initial clinical outcomes from a prospective trial," *International Journal of Radiation Oncology Biology Physics* **83**, 986-993 (2012).
- ³⁷Seibert RM, Ramsey CR, Hines JW, Kupelian PA, Langen KM, Meeks SL, and Scaperoth DD, "A model for predicting lung cancer response to therapy," *International Journal of Radiation Oncology Biology Physics* **67**, 601-609 (2007).
- ³⁸Takao S, Tadano S, Taguchi H, and Shirato H, "Computer simulation of radiotherapy for malignant tumor: A mechanical analogy method," *Journal of Biomechanical Science and Engineering* **4**, 576-588 (2009).
- ³⁹Kamrani AK and Azimi M, "Geometrical modeling of H&N cancer tumor," *Rapid Prototyping Journal* **17**, 55-63 (2011).
- ⁴⁰Azimi M, "Predictive modelling of head and neck cancer tumour based on statistical analysis," *International Journal of Collaborative Enterprise* **3**, 60-71 (2012).

- ⁴¹Mageras GS, Kutcher GJ, Leibel SA, Zelefsky MJ, Melian E, Mohan R, and Fuks Z, "A method of incorporating organ motion uncertainties into three-dimensional conformal treatment plans," *International Journal of Radiation Oncology Biology Physics* **35**, 333-342 (1996).
- ⁴²Hoogeman MS, van Herk M, Yan D, Boersma LJ, Koper PCM, and Lebesque JV, "A model to simulate day-to-day variations in rectum shape," *International Journal of Radiation Oncology Biology Physics* **54**, 615-625 (2002).
- ⁴³Zeng GG, Breen SL, Bayley A, White E, Keller H, Dawson L, and Jaffray DA, "A method to analyze the cord geometrical uncertainties during head and neck radiation therapy using cone beam CT," *Radiotherapy and Oncology* **90**, 228-230 (2009).
- ⁴⁴Bagci U, Yao JH, Miller-Jaster K, Chen XJ, and Mollura DJ, "Predicting future morphological changes of lesions from radiotracer uptake in 18F-FDG-PET images," *PLOS ONE* **8** (2013).
- ⁴⁵Schwartz LH, Colville JAC, Ginsberg MS, Wang L, Mazumdar M, Kalaigian J, Hricak H, Ilson D, and Schwartz GK, "Measuring tumor response and shape change on CT: esophageal cancer as a paradigm," *Annals of Oncology* **17**, 1018-1023 (2006).
- ⁴⁶Iyer N, Jayanti S, Lou K, Kalyanaraman Y, and Ramani K, "Three-dimensional shape searching: State-of-the-art review and future trends," *Computer-Aided Design* **37**, 509-530 (2005).
- ⁴⁷Tangelder JH and Veltkamp RC, "A survey of content based 3D shape retrieval methods," *Multimedia Tools and Applications* **39**, 441-471 (2008).

- ⁴⁸Davatzikos C, Shen DG, Mohamed A, and Kyriacou SK, "A framework for predictive modeling of anatomical deformations," *IEEE Transactions on Medical Imaging* **20**, 836-843 (2001).
- ⁴⁹Styner M, Lieberman JA, McClure RK, Weinberger DR, Jones DW, and Gerig G, "Morphometric analysis of lateral ventricles in schizophrenia and healthy controls regarding genetic and disease-specific factors," *Proceedings of the National Academy of Sciences of the United States of America* **102**, 4872-4877 (2005).
- ⁵⁰Chou YH, Tiu CM, Hung GS, Wu SC, Chang TY, and Chiang HK, "Stepwise logistic regression analysis of tumor contour features for breast ultrasound diagnosis," *Ultrasound in Medicine and Biology* **27**, 1493-1498 (2001).
- ⁵¹Tahmasbi A, Saki F, and Shokouhi SB, "Classification of benign and malignant masses based on Zernike moments," *Computers in Biology and Medicine* **41**, 726-735 (2011).
- ⁵²Yan D, Jaffray DA, and Wong JW, "A model to accumulate fractionated dose in a deforming organ," *International Journal of Radiation Oncology Biology Physics* **44**, 665-675 (1999).
- ⁵³Takao S, Tadano S, Taguchi H, Yasuda K, Onimaru R, Ishikawa M, Bengua G, Suzuki R, and Shirato H, "Accurate analysis of the change in volume, location, and shape of metastatic cervical lymph nodes during radiotherapy," *International Journal of Radiation Oncology Biology Physics* **81**, 871-879 (2011).
- ⁵⁴van der Wielen GJ, Mutanga TF, Incrocci L, Kirkels WJ, Vasquez Osorio EM, Hoogeman MS, Heijmen BJM, and de Boer HCJ, "Deformation of prostate and seminal vesicles

- relative to intraprostatic fiducial markers," *International Journal of Radiation Oncology Biology Physics* **72**, 1604-1611 (2008).
- ⁵⁵Fontenla E, Pelizzari CA, Roeske JC, and Chen GTY, "Using serial imaging data to model variabilities in organ position and shape during radiotherapy," *Physics in Medicine and Biology* **46**, 2317-2336 (2001).
- ⁵⁶Sohn M, Birkner M, Yan D, and Alber M, "Modelling individual geometric variation based on dominant eigenmodes of organ deformation: Implementation and evaluation," *Physics in Medicine and Biology* **50**, 5893-5908 (2005).
- ⁵⁷Vasquez Osorio EM, Hoogeman MS, Al-Mamgani A, Teguh DN, Levendag PC, and Heijmen BJM, "Local anatomic changes in parotid and submandibular glands during radiotherapy for oropharynx cancer and correlation with dose, studied in detail with nonrigid registration," *International Journal of Radiation Oncology Biology Physics* **70**, 875-882 (2008).
- ⁵⁸Jeong Y, Radke RJ, and Lovelock DM, "Bilinear models for inter- and intra-patient variation of the prostate," *Physics in Medicine and Biology* **55**, 3725-3739 (2010).
- ⁵⁹Sohn M, Sobotta B, and Alber M, "Dosimetric treatment course simulation based on a statistical model of deformable organ motion," *Physics in Medicine and Biology* **57**, 3693-3709 (2012).
- ⁶⁰Court L, Rosen I, Mohan R, and Dong L, "Evaluation of mechanical precision and alignment uncertainties for an integrated CT/LINAC system," *Medical Physics* **30**, 1198-1210 (2003).
- ⁶¹Thirion JP, "Image matching as a diffusion process: An analogy with Maxwell's demons," *Medical Image Analysis* **2**, 243-60 (1998).

- ⁶²Wang H, Dong L, Lii MF, Lee AL, de Crevoisier R, Mohan R, Cox JD, Kuban DA, and Cheung R, "Implementation and validation of a three-dimensional deformable registration algorithm for targeted prostate cancer radiotherapy," *International Journal of Radiation Oncology Biology Physics* **61**, 725-735 (2005).
- ⁶³Wang H, Dong L, O'Daniel J, Mohan R, Garden AS, Ang KK, Kuban DA, Bonnen M, Chang JY, and Cheung R, "Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy," *Physics in Medicine and Biology* **50**, 2887-2905 (2005).
- ⁶⁴Brock KK, "Results of a Multi-Institution Deformable Registration Accuracy Study (MIDRAS)," *International Journal of Radiation Oncology Biology Physics* **76**, 583-596 (2010).
- ⁶⁵Lele SR and Richtsmeier JT. *An invariant approach to statistical analysis of shapes*. Chapman & Hall/CRC, Boca Raton, FL 2001.
- ⁶⁶Bander M and Itzykson C, "Group theory and hydrogen atom 1," *Reviews of Modern Physics* **38**, 330-345 (1966).
- ⁶⁷Bander M and Itzykson C, "Group theory and hydrogen atom 2," *Reviews of Modern Physics* **38**, 346-358 (1966).
- ⁶⁸Morris RJ, Najmanovich RJ, Kahraman A, and Thornton JM, "Real spherical harmonic expansion coefficients as 3D shape descriptors for protein binding pocket and ligand comparisons," *Bioinformatics* **21**, 2347-2355 (2005).
- ⁶⁹Cabral B, Max N, and Springmeyer R, "Bidirectional reflection functions from surface bump maps, Proceedings of SIGGRAPH 1987," *Computer Graphics* **21**, 273-281 (1987).

- ⁷⁰Sillion FX, Arvo JR, Westin SH, and Greenberg DP, "A global illumination solution for general reflectance distributions," *Computer Graphics* **25**, 187-196 (1991).
- ⁷¹Konopliv AS, Banerdt WB, and Sjogren WL, "Venus gravity: 180th degree and order model," *Icarus* **139**, 3-18 (1999).
- ⁷²Lemoine FG, Smith DE, Rowlands DD, Zuber MT, Neumann GA, Chinn DS, and Pavlis DE, "An improved solution of the gravity field of Mars (GMM-2B) from Mars Global Surveyor," *Journal of Geophysical Research-Planets* **106**, 23359-23376 (2001).
- ⁷³Tapley B, Ries J, Bettadpur S, Chambers D, Cheng M, Condi F, Gunter B, Kang Z, Nagel P, Pastor R, Pekker T, Poole S, and Wang F, "GGM02 - An improved Earth gravity field model from GRACE," *Journal of Geodesy* **79**, 467-478 (2005).
- ⁷⁴Olsen N, Luhr H, Sabaka TJ, Mande A, Rother M, Tofner-Clausen L, and Choi S, "CHAOS-A model of the Earth's magnetic field derived from CHAMP, Orsted, and SAC-C magnetic satellite data," *Geophysical Journal International* **166**, 67-75 (2006).
- ⁷⁵Koamionkowski M, Kosowsky A, and Stebbins A, "Statistics of cosmic microwave background polarization," *Physical Review D* **55**, 7368-7388 (1997).
- ⁷⁶Yang J, Zhang D, Frangi AF, and Yang JY, "Two-dimensional PCA: A new approach to appearance-based face representation and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26**, 131-137 (2004).
- ⁷⁷Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tan PF, Westra WH, Chung CH, Jordan RC, Lu C, Kim H, Axelrod R, Silverman CC, Redmond KP, and Gillison ML, "Human papillomavirus and survival of patients with oropharyngeal cancer," *The New England Journal of Medicine* **363**, 24-35 (2010).

VITA

Adam Douglas Yock was born in Minneapolis, Minnesota on May 31, 1985 to Douglas Harold Yock, Jr., and Doris Madson Yock. He attended Breck School in Golden Valley, Minnesota from 1st through 12th grade. In May 2008, Adam received his Bachelor of Arts degree from Harvard University having concentrated in Chemistry and Physics. For the following year, he worked as a physics research fellow in the Radiation Oncology Department at the Massachusetts General Hospital. After a three month interlude exploring the Mongolian steppe, he entered the Medical Physics Ph.D. program at The University of Texas Graduate School of Biomedical Sciences with The University of Texas MD Anderson Cancer Center in 2009. There he conducted his dissertation research under the guidance of Lei Dong, Ph.D., and Laurence Court, Ph.D.

Permanent address:

11115 E De La O Road
Scottsdale, AZ 85255