University College London

Computational Modelling for Biomedical Imaging - Individual Report

---

Author:
Răzvan Valentin Marinescu
razvan.marinescu.14@ucl.ac.uk

EPSRC Centre for Doctoral Training in Medical Imaging
University College London

April 26, 2015

Introduction

Radiation therapy uses ionizing radiation to control or kill cancerous cells. It works by dmaging the DNA of the cancerous tissue leading to cellular death. In order to spare normal tissue, radiation beams are aimed from various angles and intersect at the tumour, providing a much larger dose than in the surrounding tissue. However, it is sometimes not straightforward to localise the turmour, especially due to respiratory motion. Over time, several solutions have been proposed to deal with the problem or lung motion. The simplest approach is for the patient to hold her/his breath, but this limits the acquisition time to less than 30 seconds. Implanted markers could also be used to track the tumour [1], but this is an invasive procedure that carries a certain ammount of risk. An alternative solution is to use correspondence models, which try to predict the motion of the lung based on input from a surrogate signal, which is easy to acquire during the RT intervention. The correspondence model is first trained using CT scans and the surrogate signal, and then during the RT intervention it can predict the lung position using only the surrogate signal. A lot of research has previously been done in respiratory motion modelling using different correspondence models (linear [2], polynomial [3], B-spline [4]), imaging modalities (MR [5], CT [3]) or types of surrogate signals(single signal [2], respiratory phase [6]). The aim of the project is to evaluate how well polynomial and b-spline correspondence models can model the lung deformation.

Methods

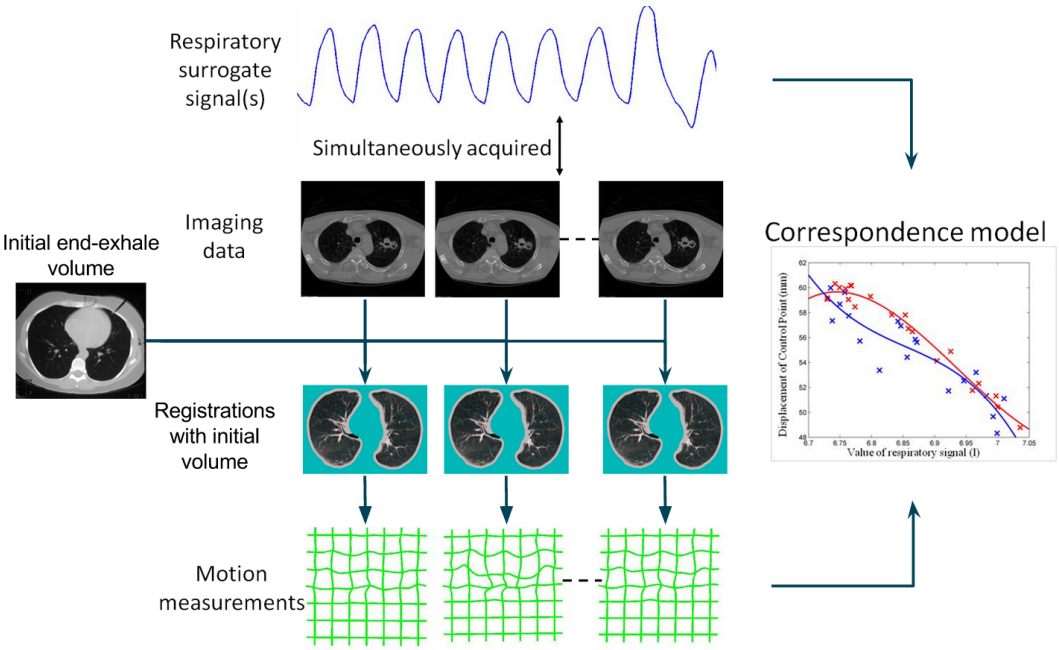


Figure 1: Diagram describing the motion model system. The respiratory surrogate signal and the Cine CT scans are acquired simoulatenously, and then these are registered with the initial end-exhale volume. The deformation grids that result from the registration and the surrogate signals are then used to build the correspondence motion model.

The imaging data given to us contained 40 Cine CT volumes acquired over 20 seconds at 5 different couch positions. Moreover, the initial end-exhale volume as well as registrations between this initial volume and the Cine CT volumes were provided. The surrogate signal was computed using the Vision RT system, and measures the volume between the chest surface and the couch. We were also given the gradient and the respiratory phase. We used this data to evaluate 9 different algorithms:

1. linear
2. quadratic
3. cubic
4. linear - phase separation
5. quadratic - phase separation
6. cubic - phase separation
7. B-spline
8. 2D linear
9. 2D quadratic

I fit the models for one voxel by solving the following GLM: $D = SC$ where D is the displacement along one direction, S is the matrix of surrogate signals and C are the coefficients. I chose to solve the GLM using the pseudo-inverse of $S^\top S$, giving $C = pinv(S^\top S)S^\top D$. I used the MATLAB function `pinv` because it can better handle singularities than the function `inv`. These 9 models have been fit for all CP displacements using a fast (vectorised) method where all the CP displacements are flattened into a huge matrix.

After all the model parameters have been calculated, a leave-one-out approach is used to estimate the CP displacements, which are then used to evaluate the models using 4 different methods: a visual assessment, deformation field error, landmark error and AIC/BIC scores. The next section presents the results of the registration assessment, model fitting, model evaluation and parameter uncertainty.

Results

Task 1 - Registration assessment

Figure 2 shows the registration results for a few representative slices and volumes over all couch positions. Both registrations perform equally well in subfigures (2a - 2e), but yielded a bad registration for couch 1, volume 6, slice 76 in subfigure 2f. Similar bad registrations have been observed in the last slice of each volume, suggesting that the registration does not always work. The images show us that there is a considerable ammount of respiratory motion across different couch position and slices which needs to be modelled. The shape of the lung is also dramatically different across different couch positions, reaching a concave shape at couch 5. Nevertheless, the segmentation and registration worked fine even for these nonstandard slices.

Task 2 & 3 - Fitting the models

The fit of the 9 correspondence models is plotted in figure 3 only for the given voxels for all the couch positions (one voxel per row). The linear, quadratic and cubic models (plots in the first column) give a reasonably good fit, especially for couch positions 4 and 5. The quadratic and cubic models don't differ much from the linear model, especially for voxels in couch positions 4 and 5. The phase separation models (middle column) seem to fit outlier points better, but at the expense of extra parameters in the model. The B-spline and the 2D linear and quadratic models seem to fit the data best, but they run the risk of overfitting. The B-spline model seems to have a very oscilatory behaviour in the first few volumes, most evident in couch 2. In conclusion, all models fit these particular voxels well, and as we expected the more complex models give a better fit than the simpler models due to the fact that they have more parameters. All these models have been fit to every voxel in every couch position and volume, and the next section presents their evaluation using visual assessment, deformation field error, landmark error and AIC/BIC scores.

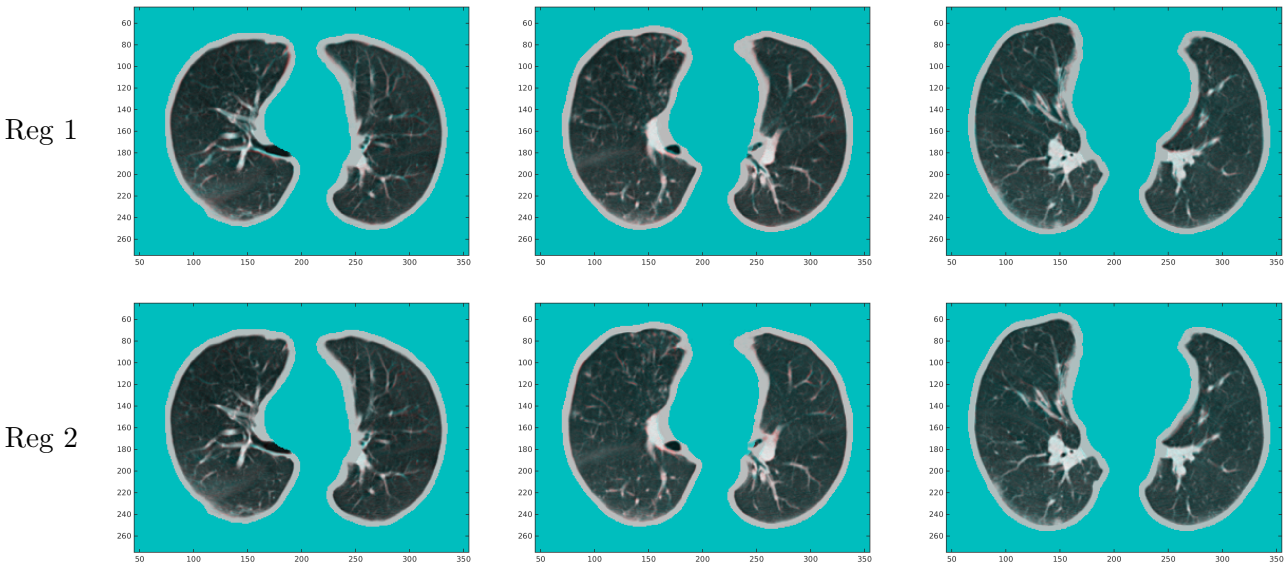
Task 4 - Evaluation of the motion models

Visual Assessment

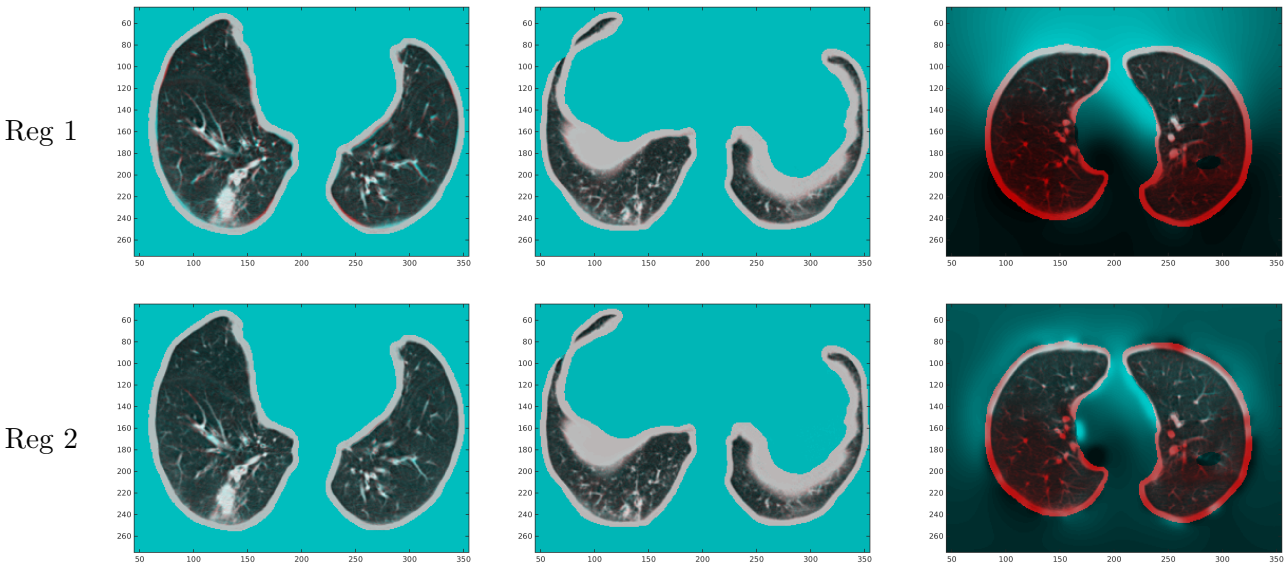
Figure 4 shows the visual assessment for all the 9 models for couch 1, volume 1, slice 66. There is considerably more motion compared to the registration assessment in task 1. The results are quite good, with essentially no visible differences between the estimated deformation across the models. The motion of the outer shape of the lung has been almost perfectly estimated, with tiny errors only at the top corner of the right lung. More deformation errors are found at the vasculature of the right lung. We therefore conclude that all 9 models perform equally well.

Deformation Field Error

Figure 5 shows the deformation field error for the 10 Cine CT volumes across all 5 couch positions (note that we couldn't show results for the other 30 Cine CT scans because we weren't given the required data). The left column shows the mean error, while the right column shows the standard deviation. The ammount of motion varies slightly across the Cine CT scans but is usually less than 3mm. Similarly,

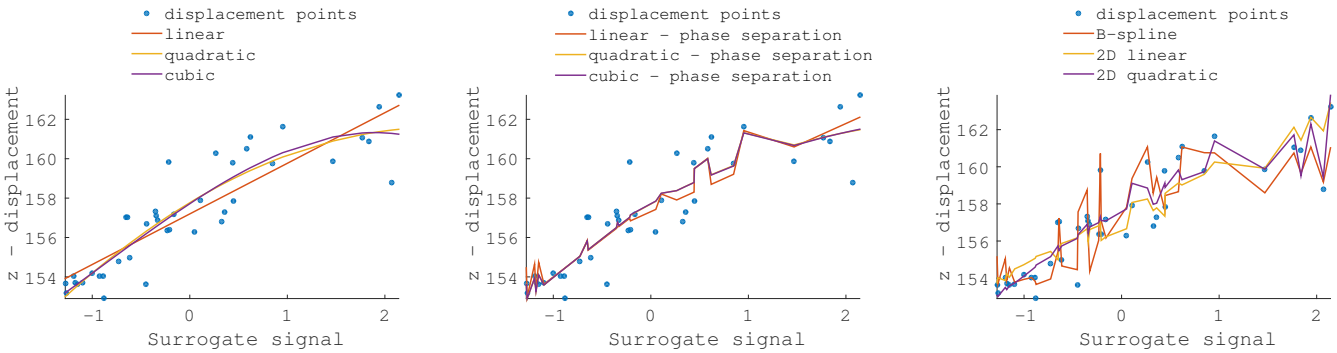


(a) Couch 1, volume 1, slice 66 (b) Couch 2, volume 5, slice 60 (c) Couch 3, volume 2, slice 49

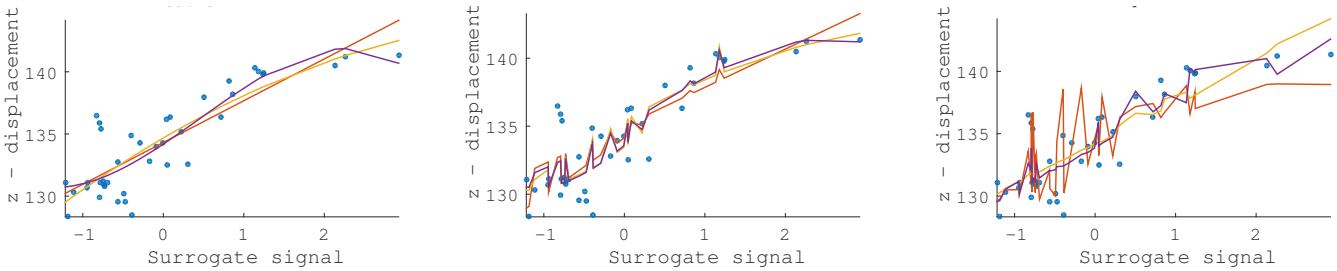


(d) Couch 4, volume 8, slice 37 (e) Couch 5, volume 4, slice 25 (f) Couch 1, volume 6, slice 76

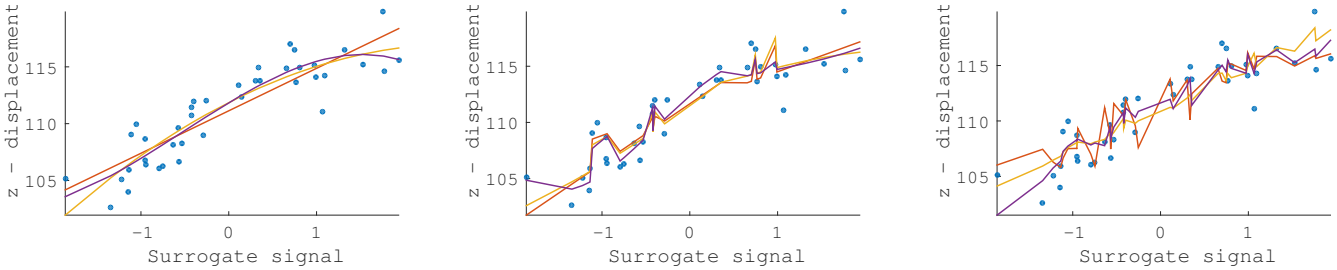
Figure 2: Registration results at several different slices and volumes for each couch position. In each subfigure, registration 1 and 2 are shown in the upper and lower figures respectively. Both registrations worked fine in these examples apart from the slice in subfigure 2f.



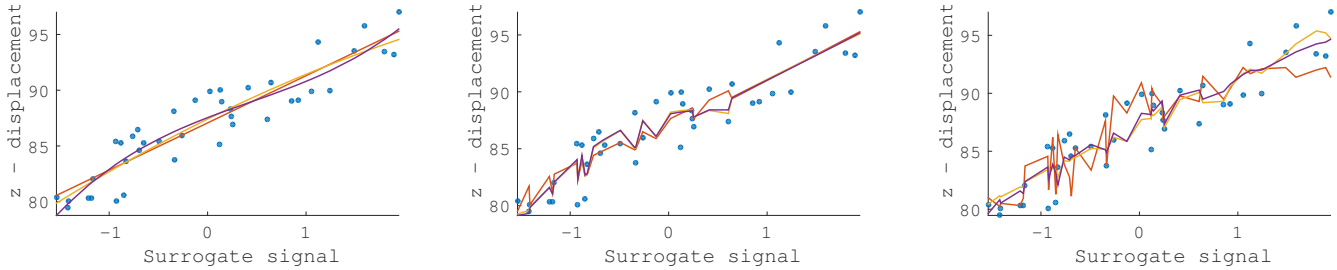
(a) Fit for couch position 1, voxel (30,40,33,1,3)



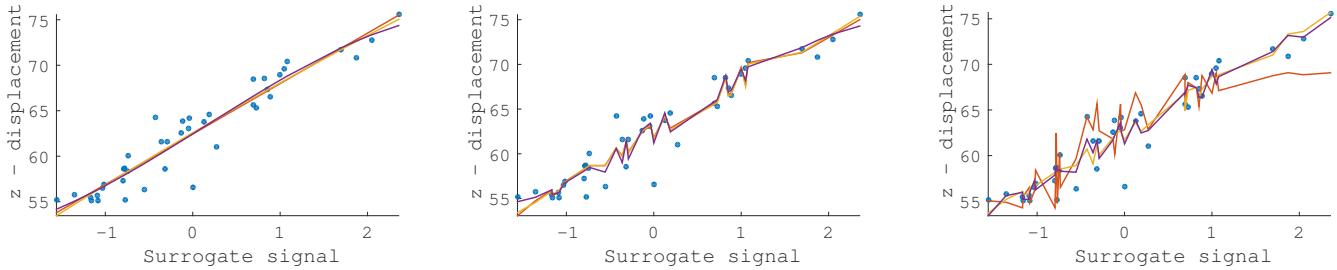
(b) Fit for couch position 2, voxel (30,40,28,1,3)



(c) Fit for couch position 3, voxel (30,40,23,1,3)

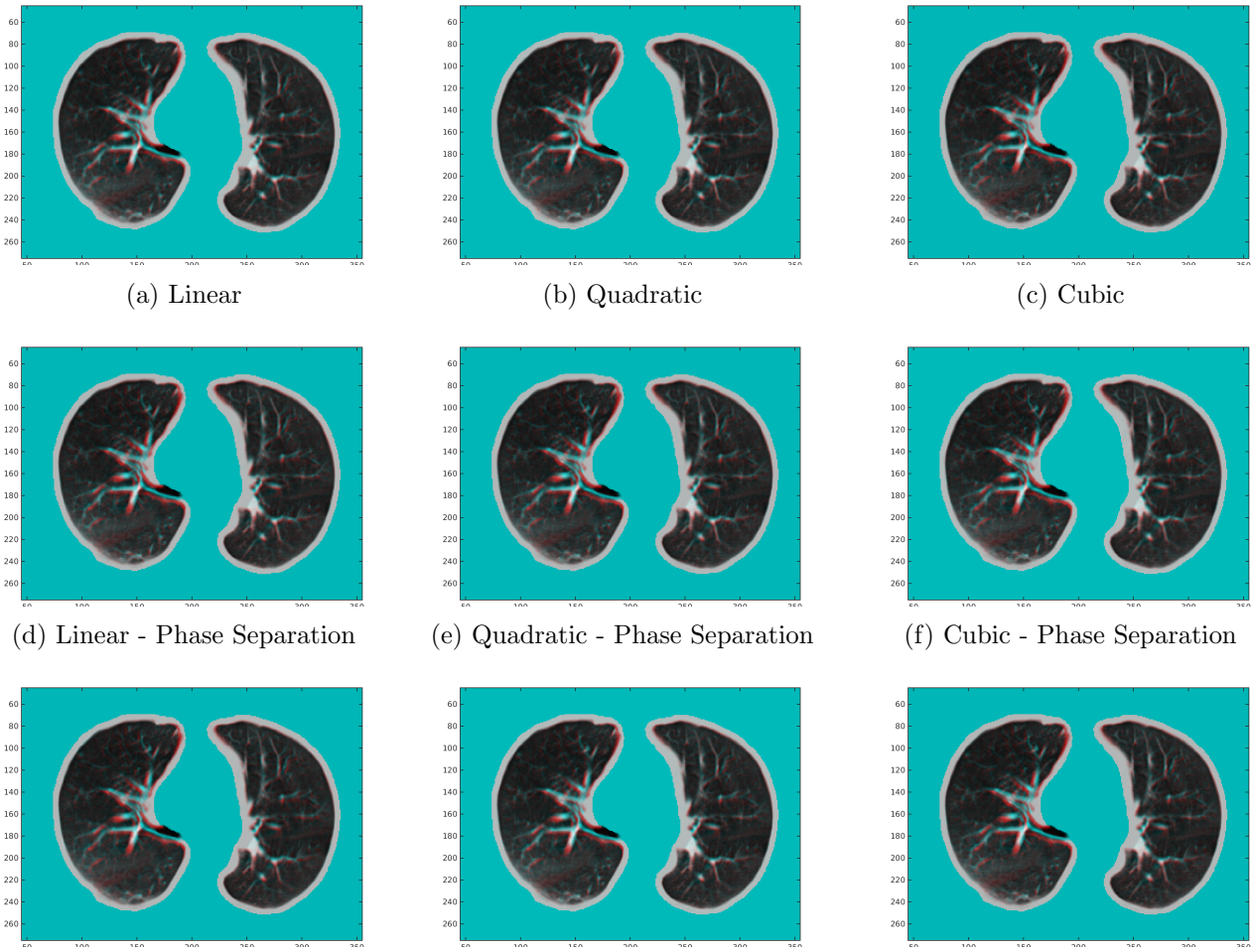


(d) Fit for couch position 4, voxel (30,45,18,1,3)



(e) Fit for couch position 5, voxel (30,45,13,1,3)

Figure 3: Fit of 9 different models across 5 voxels each from a different couch position. The first column shows the fit for the three basic models: linear, quadratic and cubic. The second column shows the same models with phase separation, while the last column shows the fit for the B-spline, 2D linear and 2D quadratic models.



(a) Linear (b) Quadratic (c) Cubic

(d) Linear - Phase Separation (e) Quadratic - Phase Separation (f) Cubic - Phase Separation

(g) B-spline (h) Linear 2D (i) Quadratic 2D

Figure 4: Visual assessment of the 9 deformation models using leave-one out cross-validation on couch 1, volume 1, slice 66. All models perform equally well, with most errors at the vasculature and at the top of the right lung.

standard deviation is usually below 1.2. The B-spline model seems to have larger error means and standard deviations especially in couch positions 3, 4 and 5. This might be due to overfit of the data. The other models have similar errors across all couch positions. We therefore conclude that all models apart from the B-spline model perform equally well.

Landmark Error

Figure 6 shows the landmark errors for all the 9 models across the 5 couch positions, this time using all 40 Cine CT volumes. Moreover, the landmark error using the original registration is also plotted in black. The landmark error is usually between 1 and 4mm. All models perform equally good, with no clear winner from the plots. In particular, the cubic - phase separation model seems to have a spike in landmark error at couch position 2, volume 24, which might be due to overfitting. It should be noted that for several Cine CT volumes, the landmark error of some models is lower than the landmark error of the original registration. In the couch 5 plot, missing values arrise due to the fact that the landmark points could not be located in those Cine CT volumes as there was too much motion.

Model Ranking

Finally, in order to rank the 9 different correspondence models, AIC and BIC scores have been calculated:

Model	AIC	BIC	Mean Deformation Error (mm)
cubic	90841746	90841805	1.161
cubic - phase separation	91489516	91489634	1.263
quadratic	92590097	92590141	1.169
quadratic - phase separation	93364438	93364527	1.207
linear	94080458	94080487	1.175
linear - phase separation	94325216	94325275	1.198
2D quadratic	96879272	96879360	1.240
2D linear	97780958	97781002	1.182
B-spline	118933267	118933326	1.706

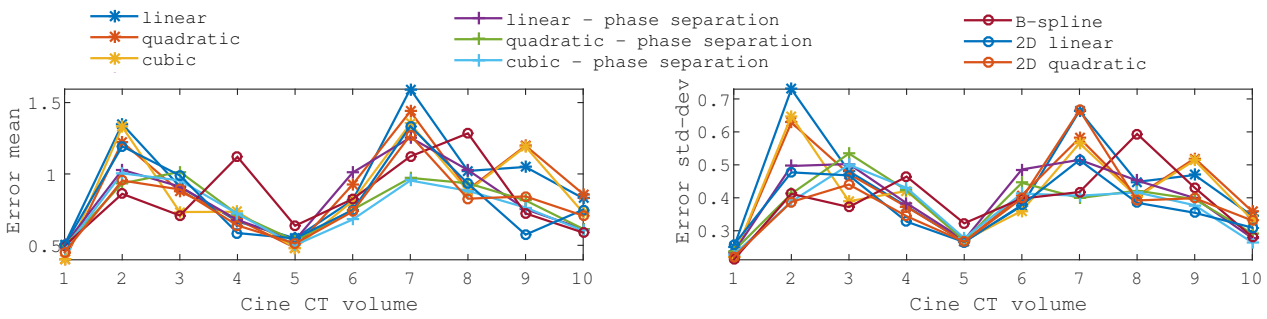
The model ranking (ordered by AIC score) shows that the cubic model has the lowest AIC/BIC score, followed by cubic-phase separation, quadratic and quadratic-phase separation. At the bottom end of the table, the B-spline, 2D linear and 2D quadratic models have the highest AIC/BIC scores. These results agree with the previous ones obtained using the deformation field error and landmark error. Mean deformation error for all models ranged from 1.16 mm (cubic model) to 1.70 mm(B-spline model).

The advantage of using all these different methods of assessing the models is that it can increase our confidence in the results. Furthermore, each methods assesses the models from different angles, so one method of assessment could potentially show some interesting insights that the others cannot offer. The visual assessment offers a broad overview of the prediction accuracy and help identify in which are the areas of the lungs motion is not properly predicted. The assessment using the deformation field error and AIC/BIC scores is useful for finding out which models perform best and even rank them, while the landmark error assessment is unbiased as it uses external input from a clinician. The leave one out method was used in order to provide a more unbiased prediction of the deformation by not using the same data for both training and testing. As there were only 40 volumes available and in order to have as large a training sample as possible, 39 of the datapoints were used for training and 1 for testing. Nevertheless, some problems might arrise with the leave one out approach, such as having highly correlated samples, especially those that are very close in time.

Task 5 - Parameter uncertainty

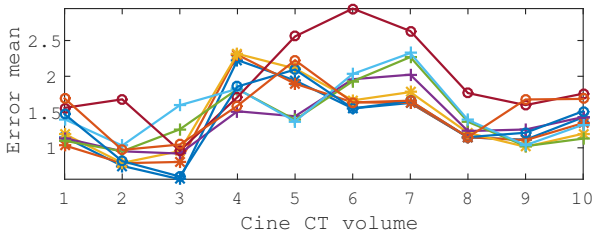
We estimated parameter uncertainty in the 3 basic models using 4 different methods. The reason why these were chosen is explained next to them:

- **Parametric Bootstrap**: used because it assumes noise model is correct
- **Residual Bootstrap**: uses residual values instead of random samples from a noise model

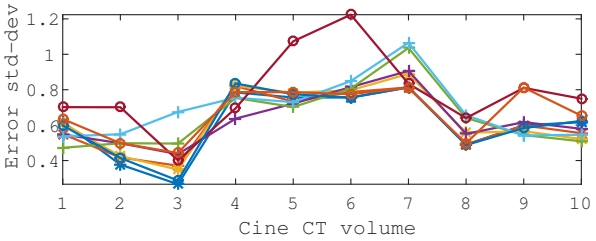


(a) Couch 1 - Mean Def. Error

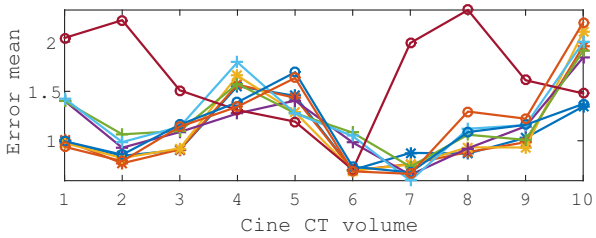
(b) Couch 1 - Std. Dev. of Def. Error



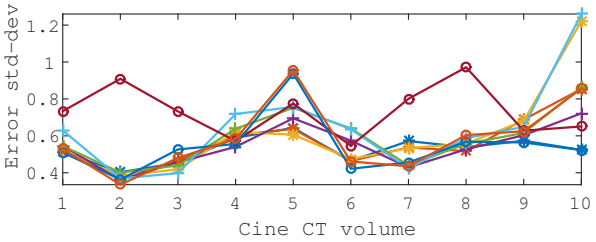
(c) Couch 2 - Mean Def. Error



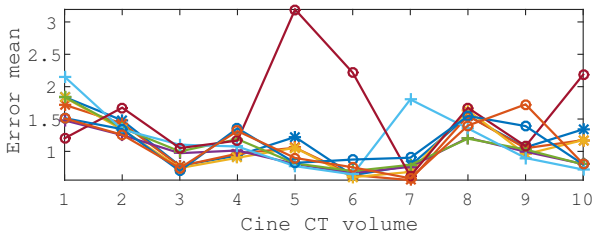
(d) Couch 2 - Std. Dev. of Def. Error



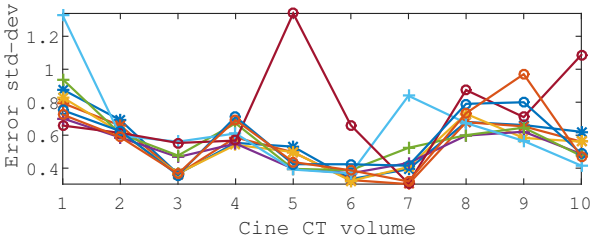
(e) Couch 3 - Mean Def. Error



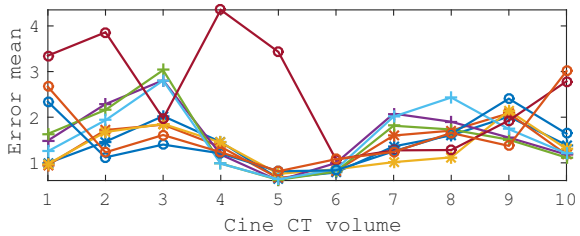
(f) Couch 3 - Std. Dev. of Def. Error



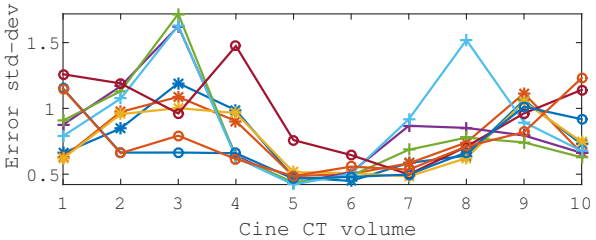
(g) Couch 4 - Mean Def. Error



(h) Couch 4 - Std. Dev. of Def. Error

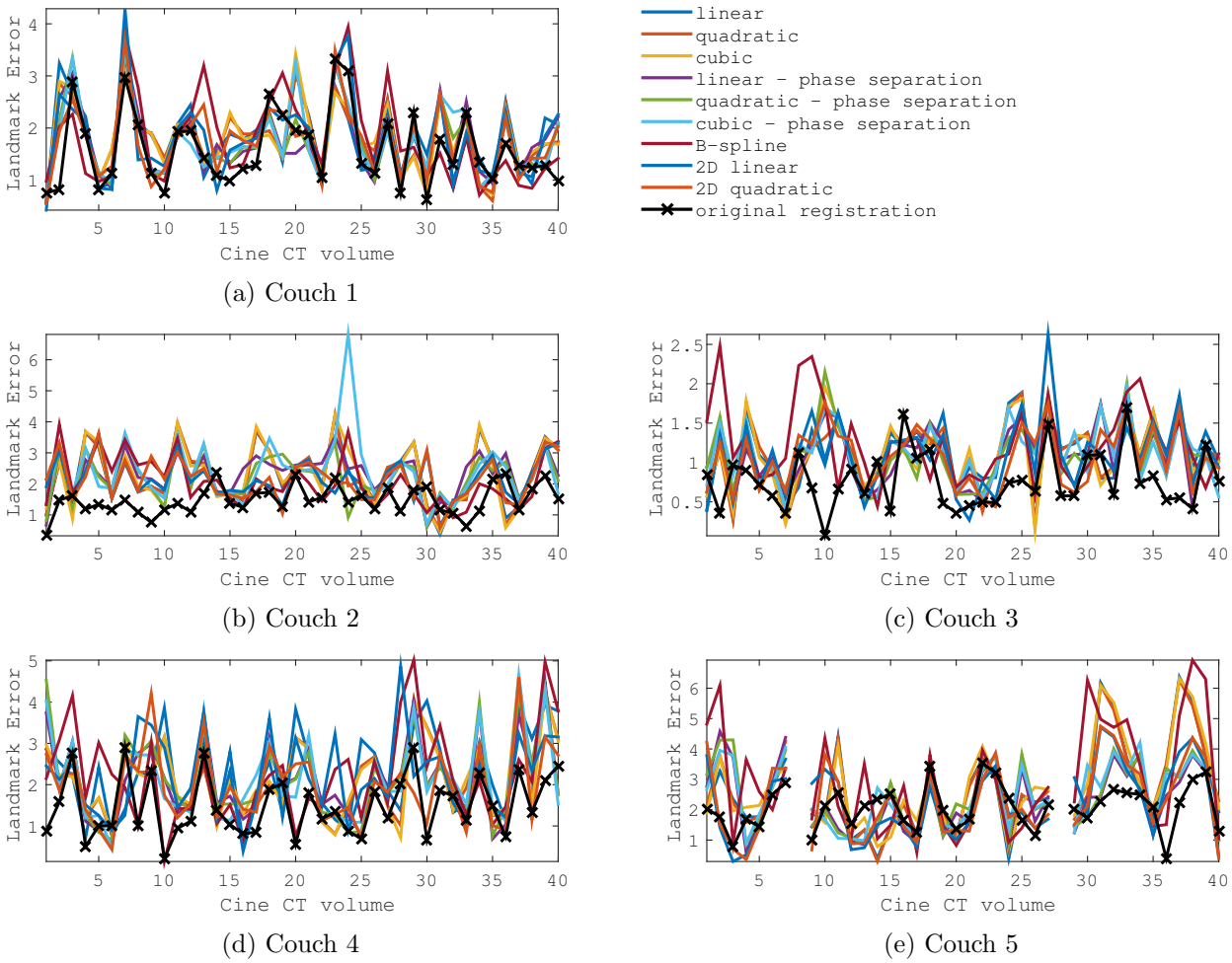


(i) Couch 5 - Mean Def. Error



(j) Couch 5 - Std. Dev. of Def. Error

Figure 5: Deformation Error mean and standard deviation for all 5 couch positions, 10 Cine CT volumes and 9 models. All models seem to perform equally well, apart from the B-spline model, which has some large error means and standard deviations especially in couch positions 3,4 and 5. This might be due to overfit of the data.



(a) Couch 1

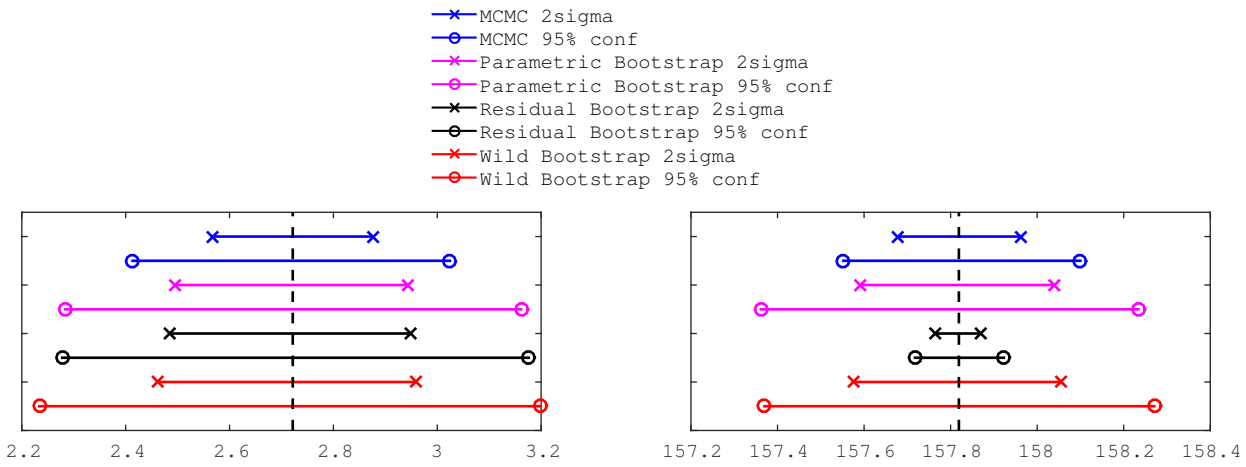
(b) Couch 2

(c) Couch 3

(d) Couch 4

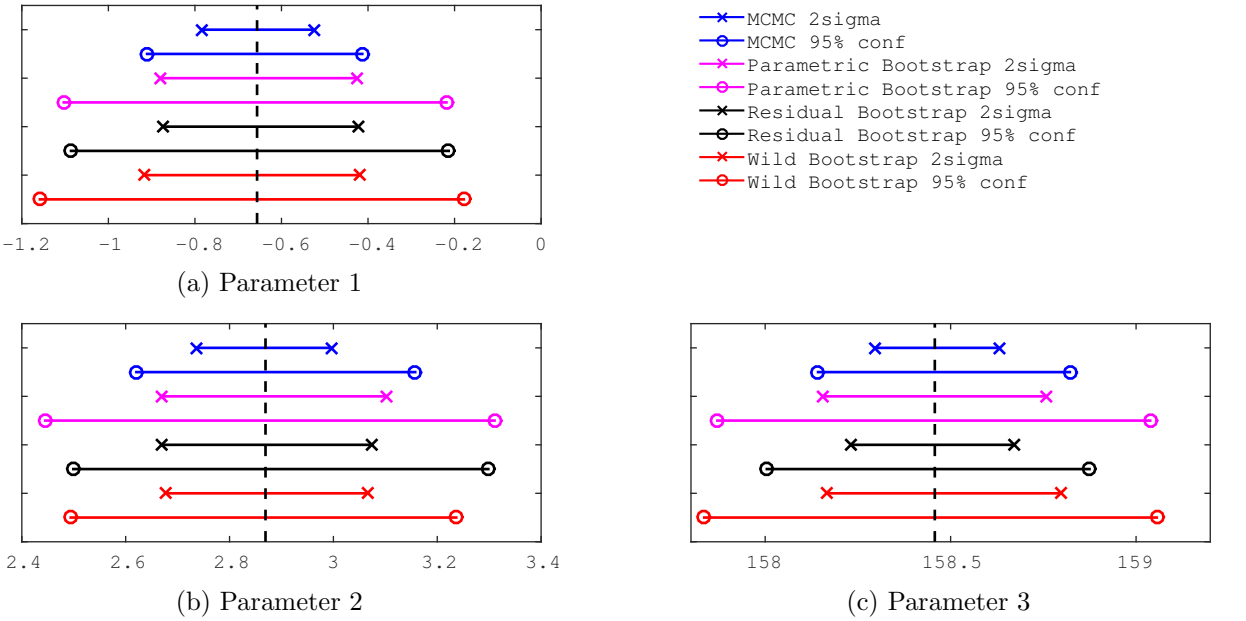
(e) Couch 5

Figure 6: Landmark error for all 5 couch positions, 40 Cine CT volumes and 9 models as well as for the original registration (in black). All models perform equally well and sometimes even have a lower landmark error than the original registration. The missing values in couch 5 represent landmark points that could not be located in the Cine CT volumes as there was too much motion.



(a) Parameter 1 (b) Parameter 2

Figure 7: Parameter uncertainty for the linear model.



(a) Parameter 1

(b) Parameter 2 (c) Parameter 3

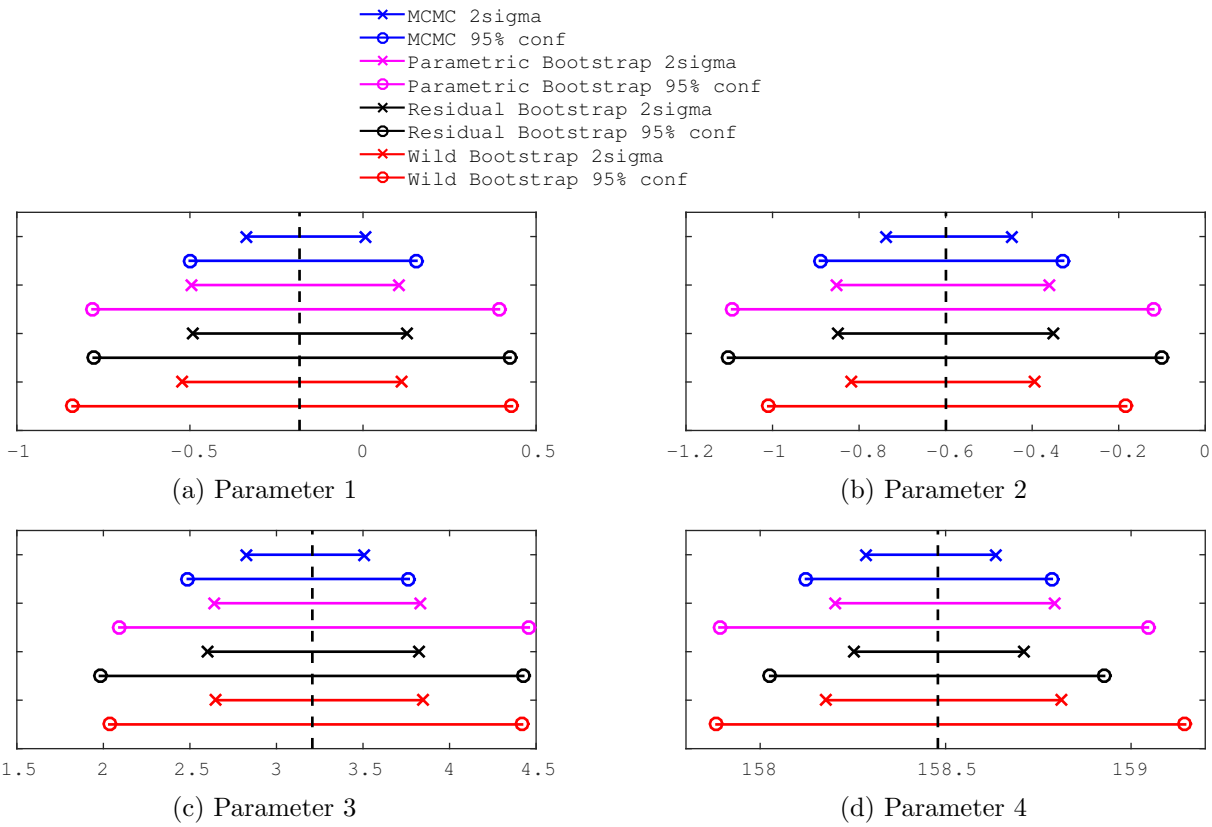
Figure 8: Parameter uncertainty for the quadratic model.

- **Wild Bootstrap**: suitable for heteroskedastic data (i.e. having populations with different variances)
- **MCMC**: used because it perturbs the parameters rather than the dataset

Results for the linear, quadratic and cubuc models are plotted in figures 7, 8 and 9 respectively. For all plots, we notice that the fitted values were always in the center of the 95% confidence intervals and 2-sigma ranges. For each model, the value of the intercept (last parameter in each plot) is always around 157.5-158 while the value of the linear coefficient (penultimate parameter in each plot) is around 2.5-4. For the quadratic and cubic models, we notice that the coefficients of the quadratic and cubic terms are much smaller (betwen -1 and 0), suggesting that there is not much extra fit that these higher-order models can give. The corresponding parameter distribution for the parametric bootstraps are plotted in appendix figures 10, 11 and 12 respectively.

Discussion

Our results show that correspondence models can be successfully used to predict lung motion, with a precision of approximatelly 2-3mm. It should be noted that part of the error also comes from the registration that didn't properly align the end-exhale Cine CT slice with the other slices. We also concluded that the



(a) Parameter 1

(b) Parameter 2

(c) Parameter 3

(d) Parameter 4

Figure 9: Parameter uncertainty for the cubic model.

cubic model is the most suitable, as it has the lowest AIC and BIC scores, followed by cubic with phase separation and quadratic. The 3 methods of evaluation (visual assessment, deformation field error and landmark error) yielded consistent results. Finally, we also calculated the ammount of uncertainty in the fitted parameters.

Bibliography

[1] Hiroki Shirato, Shinichi Shimizu, Tatsuya Kunieda, Kei Kitamura, Marcel van Herk, Kenji Kagei, Takeshi Nishioka, Seiko Hashimoto, Katsuhisa Fujita, Hidefumi Aoyama, et al. Physical aspects of a real-time tumor-tracking system for gated radiotherapy. *International Journal of Radiation Oncology* Biology* Physics*, 48(4):1187–1195, 2000.

[2] Achim Schweikard, Greg Glosser, Mohan Bodduluri, Martin J Murphy, and John R Adler. Robotic motion compensation for respiratory movement during radiosurgery. *Computer Aided Surgery*, 5(4):263–277, 2000.

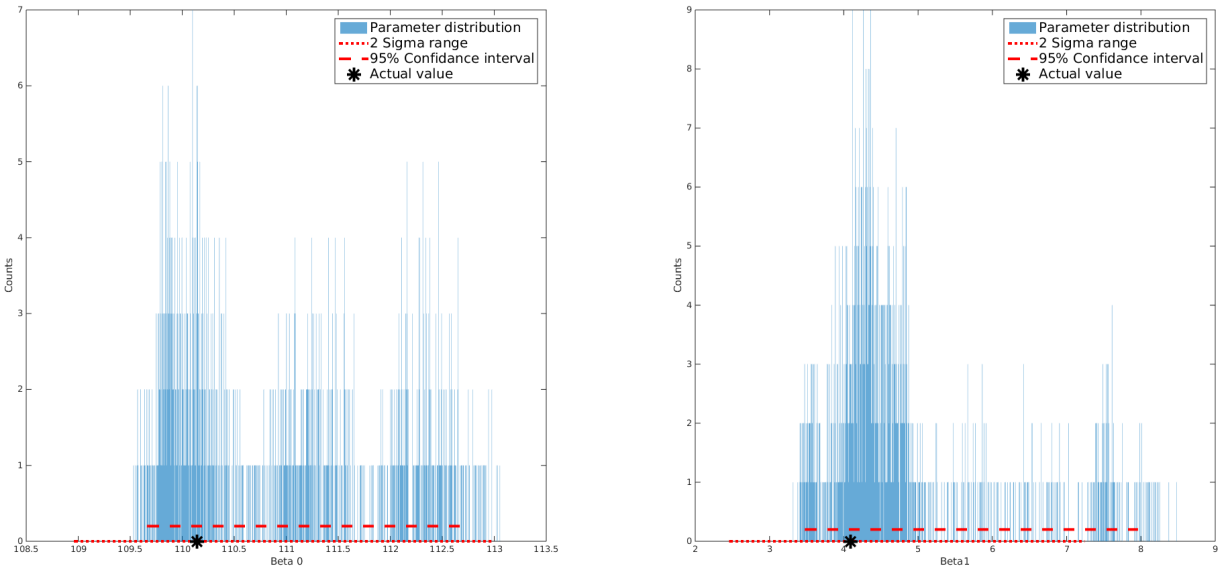
[3] JR McClelland, AG Chandler, JM Blackall, S Ahmad, DB Landau, and DJ Hawkes. 4d motion models over the respiratory cycle for use in lung cancer radiotherapy planning. In *Medical Imaging*, pages 173–183. International Society for Optics and Photonics, 2005.

[4] Christian Buerger, RE Clough, Andrew P King, Tobias Schaeffter, and Claudia Prieto. Nonrigid motion modeling of the liver from 3-d undersampled self-gated golden-radial phase encoded mri. *Medical Imaging, IEEE Transactions on*, 31(3):805–815, 2012.

[5] Dirk Manke, Kay Nehrke, and Peter Börnert. Novel prospective respiratory motion correction approach for free-breathing coronary mr angiography using a patient-adapted affine motion model. *Magnetic Resonance in Medicine*, 50(1):122–131, 2003.

[6] Guy Shechter, Cengizhan Ozturk, Jon R Resar, and Elliot R McVeigh. Respiratory motion of the heart from free breathing coronary angiograms. *Medical Imaging, IEEE Transactions on*, 23(8):1046–1056, 2004.

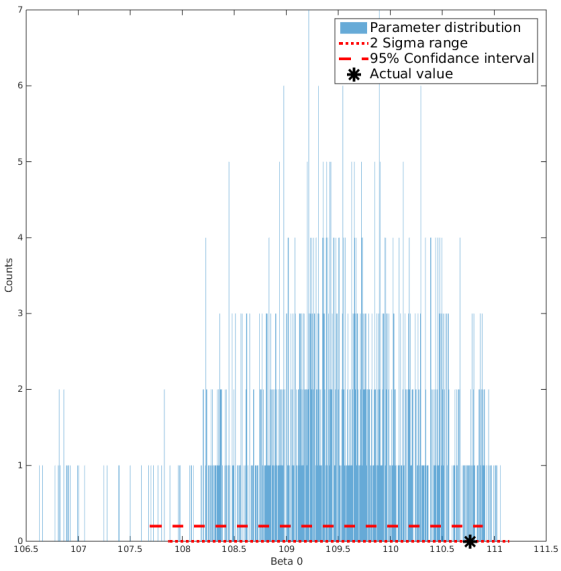
Appendix



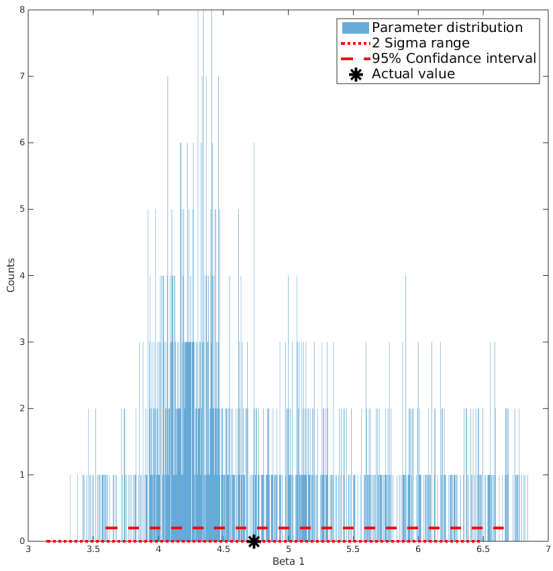
(a) Parameter 1

(b) Parameter 2

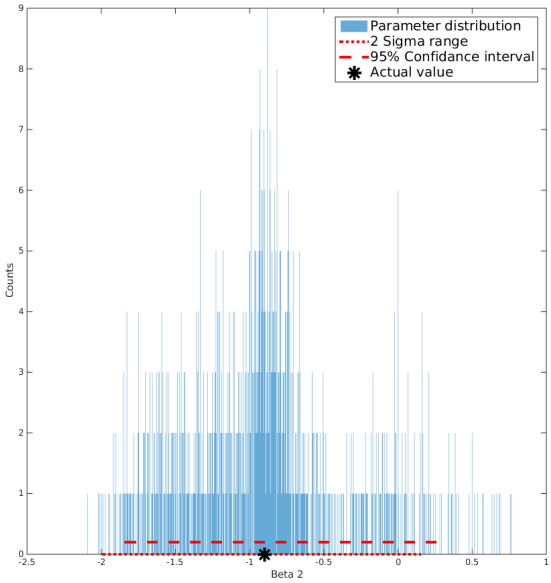
Figure 10: Parametric bootstrap for the linear model - parameter distributions.



(a) Parameter 1

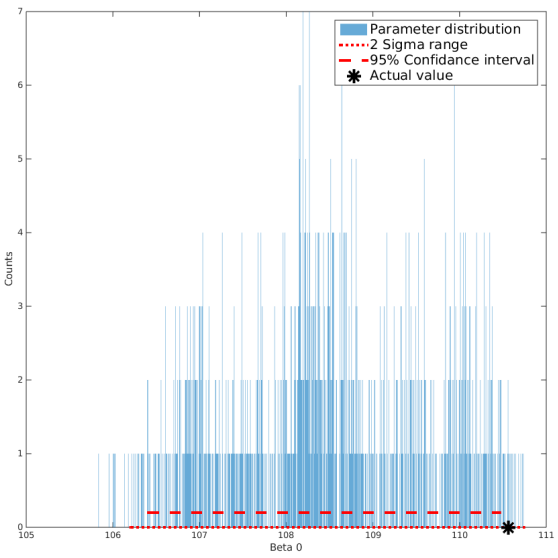


(b) Parameter 2

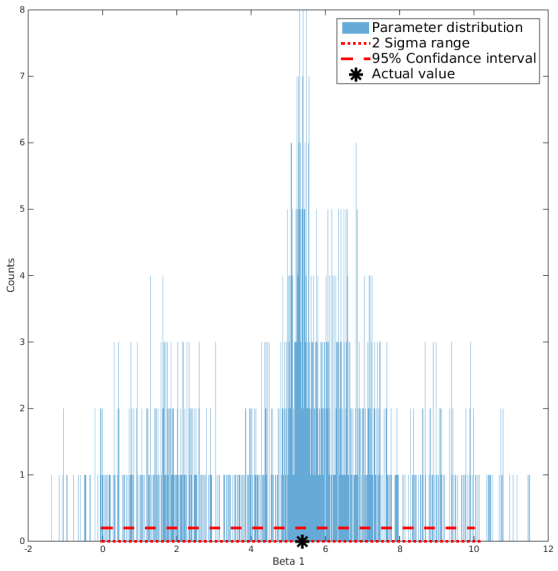


(c) Parameter 3

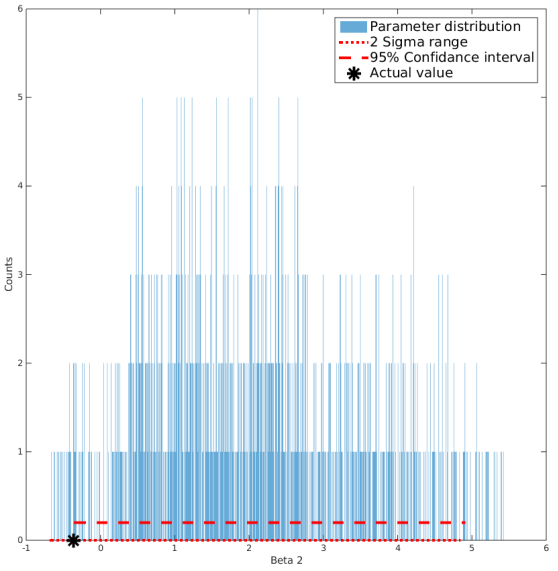
Figure 11: Parametric bootstrap for the quadratic model - parameter distributions.



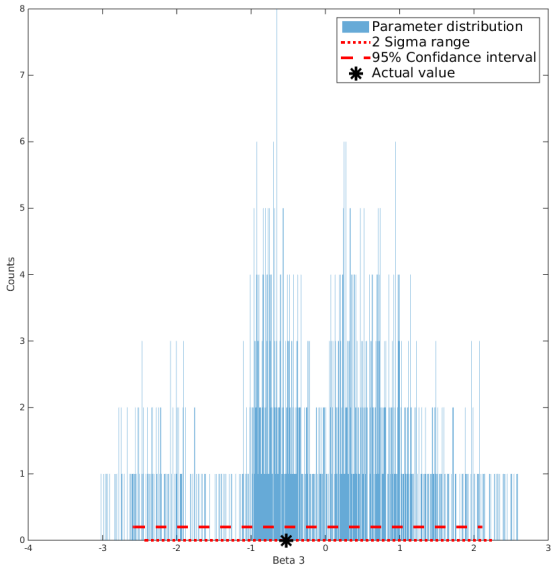
(a) Parameter 1



(b) Parameter 2



(c) Parameter 3



(d) Parameter 4

Figure 12: Parametric bootstrap for the cubic model - parameter distributions.