



Performance comparison of prediction filters for respiratory motion tracking in radiotherapy

Journal Article**Author(s):**

Jöhl, Alexander ; Ehrbar, Stefanie; Guckenberger, Matthias; Klöck, Stephan; Meboldt, Mirko; Zeilinger, Melanie; Tanadini-Lang, Stephanie; Schmid Daners, Marianne 

Publication date:

2020-02

Permanent link:

<https://doi.org/10.3929/ethz-b-000380713>

Rights / license:

[In Copyright - Non-Commercial Use Permitted](#)

Originally published in:

Medical Physics 47(2), <https://doi.org/10.1002/mp.13929>

Performance comparison of prediction filters for respiratory motion tracking in radiotherapy

Alexander Jöhl

*Product Development Group Zurich, Department of Mechanical and Process Engineering, ETH Zurich, Zurich, Switzerland
Department of Radiation Oncology, University Hospital Zurich, Zurich, Switzerland*

Stefanie Ehrbar, Matthias Guckenberger, and Stephan Klöck

*Department of Radiation Oncology, University Hospital Zurich, Zurich, Switzerland
University of Zurich, Zurich, Switzerland*

Mirko Meboldt

Product Development Group Zurich, Department of Mechanical and Process Engineering, ETH Zurich, Zurich, Switzerland

Melanie Zeilinger

Institute for Dynamic Systems and Control, Department of Mechanical and Process Engineering, ETH Zurich, Zurich, Switzerland

Stephanie Tanadini-Lang

*Department of Radiation Oncology, University Hospital Zurich, Zurich, Switzerland
University of Zurich, Zurich, Switzerland*

Marianne Schmid Daners^{a)}

Product Development Group Zurich, Department of Mechanical and Process Engineering, ETH Zurich, Zurich, Switzerland

(Received 24 April 2019; revised 28 October 2019; accepted for publication 6 November 2019; published xx xxxx xxxx)

Purpose: In precision radiotherapy, the intrafractional motion causes substantial uncertainty. Traditionally, the target volume is expanded to cover the tumor in all positions. Alternative approaches are gating and adaptive tracking, which require a time delay as small as possible between the actual tumor motion and the reaction to effectively compensate the motion. Current treatment machines often exhibit large time delays. Prediction filters offer a promising means to mitigate these time delays by predicting the future respiratory motion.

Methods: A total of 18 prediction filters were implemented and their hyperparameters optimized for various time delays and noise levels. A set of 93 traces were standardized to a sampling frequency of 25 Hz and smoothed using the Fourier transform with a 3 Hz cutoff frequency. The hyperparameter optimization was carried out with ten traces, and the optimal hyperparameters were evaluated on the remaining 83 traces.

Results: For smooth traces, the wavelet least mean squares prediction filter and the linear filter reached normalized root mean square errors of below 0.05 for time delays of 160 and 480 ms, respectively. For noisy signals, the performance of the prediction filters deteriorated and led to similar results.

Conclusions: Linear methods for prediction filters are sufficient for respiratory motion signals. Reducing the measurement noise generally improves the performance of the prediction filters investigated in this study, even during breathing irregularities. © 2019 American Association of Physicists in Medicine [<https://doi.org/10.1002/mp.13929>]

Key words: motion compensation, prediction filter, respiratory motion

Abbreviations

ANFIS adaptive neuro-fuzzy interference system
 e_{nRMS} normalized root mean square error
ERLS extended recursive least squares
IMM interacting multiple model
KDE kernel density estimation
KFCA kalman filter constant acceleration
KFCV kalman filter constant velocity
LCMEKF local circular model extended Kalman filter
LF linear filter
LOESS local regression
LR linear regression

MULIN multistep linear
NERLS nonlinearly extended recursive least squares
NLMS normalized least mean squares
NN neural network
RLS recursive least squares
RVM relevance vector machines
SNR signal-to-noise ratio
SVR support vector regression
TAKENS takens prediction filter
WLMS wavelet least mean squares

1. INTRODUCTION

Modern radiation therapy techniques can focus the ionizing radiation precisely on the tumor volume. However, motion causes substantial uncertainty during treatment delivery, for instance due to respiration. It was reported that the peak-to-peak motion amplitudes of lung and liver tumors can amount to up to 38 mm¹ and 34 mm,² respectively. If this motion is not adequately considered, the effectiveness of the treatment can be considerably reduced.³

The established approach to account for intrafractional motion is to expand the target volume such that the tumor is guaranteed to be covered in all positions. This approach assures the accurate radiation dose coverage of the tumor, but it increases the irradiation dose to the healthy tissue surrounding the tumor. Several approaches were proposed to overcome the uncertainty due to intrafractional motion,⁴ which aim at reducing the dose to healthy tissue while ensuring the full dose coverage of the tumor. During gated treatment, the motion of the tumor is observed and the radiation beam is only switched on if the tumor is in a specific position.⁴ With the CyberKnife⁵ or the Vero⁶ system, the tumor motion is continuously compensated by moving the beam, while with multi-leaf collimator (MLC) tracking,⁷ the beam is modified according to the patient movement. During couch tracking, the patient is moved with the robotic couch⁸ to compensate the tumor motion.

These proposed approaches require fast action of the systems involved to be effective at reducing the uncertainty due to intrafractional motion. For example, gated treatments require short reaction times to switch off the beam. If this reaction time is long, the tumor leaves the irradiated volume before the beam is off, which leads to an unnecessary exposure of healthy tissue and a lack of dose to the tumor. Tracking approaches with a continuous compensation of the tumor motion require the systems involved to move synchronized to the tumor. In that case, not only the reaction times but also the limited dynamics of the systems become important. The limited dynamics (e.g., mass inertia) cause the systems to lag behind the tumor motion. As a consequence, the reaction times and the time lags are large sources of errors in tumor motion compensation. We refer to the sum of reaction time and time lag as time delay.

One option to mitigate this source of errors is to improve the systems such that they have faster reaction times and faster dynamics to reduce the time delays. However, this approach requires a large investment because it involves hardware changes. Another option is to compensate for the time delay by exploiting the patterns of the tumor motion to predict the motion ahead. In this approach, an algorithm, often called prediction filter, takes the measured signal of the tumor motion as an input, and outputs a prediction of the signal. The subsequent compensatory motion of the systems reacting to the new signal thus matches the actual tumor motion. Such prediction filters only require software modifications.

A large number of prediction filters are presented in literature. A wide variety of techniques for prediction filters have

been proposed, including Kalman filters, autoregressive moving average models, wavelet decomposition, fuzzy logic models, neural nets, support vector regression, nonlinear dynamics identification, and combinations thereof.^{9,10} Prediction filters have also been compared to each other: The support vector regression approach was compared to the neural network approach¹¹ and found to be superior. A larger comparison was carried out with seven prediction filters,¹² in which not only linear regression methods but also nonlinear methods such as support vector regression were compared. The support vector regression approach was found to be superior to all other evaluated techniques. In a further study,¹³ six prediction filters were compared, and the wavelet least mean squares prediction filter was found to provide the lowest prediction error. In other work,¹⁴ linear regression methods were compared to nonlinear methods such as neural networks and support vector regression. Neural networks were found to perform best. In the work of Murphy and Dieterich,¹⁵ linear regression methods were compared to the neural network approach, and the neural network was found to be superior. Furthermore, the effect of smoothing the respiratory traces before applying two different prediction filters was investigated.¹⁶ Smoothing was found to improve the performance of the prediction filters. The estimation of the performance of the prediction filter before applying them to a patient's respiratory motion was investigated as well.¹⁷ The approach was to find characteristic features of the traces and then correlate the features with the prediction filter's performance.

The comparisons of all prediction filters presented in the literature are limited in the number of prediction filters investigated and additionally, in the number of respiratory motion traces applied. In the presented study, we implemented 18 prediction filters and optimized their hyperparameters regarding respiratory motion traces, varying time delays and noise levels. Hyperparameters are parameters of the training procedure itself, which identifies the normal parameters' values. The normal parameters are then used for the prediction. The prediction filters were evaluated on a set of 93 one-dimensional traces with two time delays and three noise levels. The aim of this study is the performance comparison of these 18 different prediction filters under different noise levels and time delays. For every given noise level, time delay, and prediction filter, the hyperparameters were optimized with the same optimization algorithm on the same set of real patient data and evaluated on a different set of real patient data to avoid overfitting. The results of the prediction filters were analyzed with focus on the effect of noise level variation using real patient data. To the best of our knowledge, this study presents the largest comparison of prediction filters in the literature.

2. METHODS AND MATERIALS

2.A. Respiratory motion data

The respiratory motion traces were collected from recordings during CyberKnife treatments.¹⁸ Some of the traces

exhibited long periods of invalid measurements (marker not detected) and thus were excluded. The remaining traces contained short periods of invalid samples; therefore, every trace was resampled at a sampling rate of 25 Hz via spline interpolation. For investigating noise level variation, smooth traces were first constructed by filtering the original traces via Fourier transform, such that only frequencies below 3 Hz were contained in the smoothed traces. Each smoothed trace was then corrupted with added Gaussian noise, with signal-to-noise ratios (SNR) of 30 and 20 dB, respectively, see Fig. 1. Consequently, the noise variation in this study had three levels: smooth, SNR of 30 dB, and SNR of 20 dB. In total, 93 different respiratory motion traces were included, where the recording durations ranged from 675.44 s to 7883.64 s (median 4364.04 s). The mean respiration period, mean amplitude, and mean speed of the traces are shown in Fig. 2. Significant drift occurs in 28 of 93 traces, where a significant drift was defined as a change in baseline of more than 5 mm over a duration of 100 s. Sudden large motion (defined as an amplitude excursion of more than twice the amplitude of the three previous respiratory cycles) occurred in 67 of the 93 traces. Time delays of four (160 ms) and 12 (480 ms) time steps were investigated based on findings by Hoogeman et al.⁵ and Cho et al.¹⁹ The range of respiratory frequencies varied from 0.16 to 0.40 Hz with a median of 0.27 Hz. The implemented code for the described procedure is available in the supplementary material.

2.B. Prediction filters

In the following, all of the 18 prediction filters that were investigated are presented. The prediction filters were selected to represent a wide range of approaches that were previously published. The algorithms were implemented following the respective publications. However, in our implementations, learning is continuous, that is, the samples used for training of the prediction filters' parameters were updated at every time step by adding the current and discarding the oldest time sample. In general, the prediction filters were

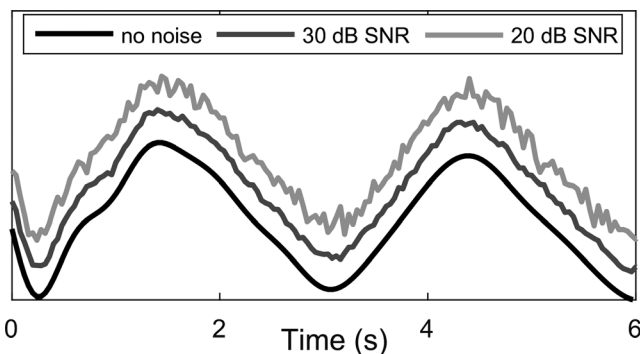


FIG. 1. Qualitative visualization of a standardized trace with the varying signal-to-noise ratios (SNRs). For each SNR, the trace is depicted in an offset to allow an easy visual comparison. The Gaussian noise was added after the respiratory motion traces were resampled to 25 Hz and smoothed using a Fourier transform with a cutoff frequency of 3 Hz.

retrained at every time step, exceptions due to computational challenges are stated in the following short descriptions of the prediction filters:

1. The adaptive neuro-fuzzy interference system (ANFIS) prediction filter is a combination of a neural network with fuzzy logic.²⁰ For this prediction filter, the Fuzzy Logic Toolbox (The MathWorks Inc., Natick, MA, USA) was used. At every time step, the ANFIS predicted the value based on the past values. To reduce computation time, the training of the ANFIS was carried out at every tenth time step.
2. The local regression (LOESS) prediction filter utilizes polynomial regression using past values of the trace, but only those that are similar (Euclidean distances below a threshold) to the current values.²¹
3. The normalized least mean squares (NLMS) prediction filter uses a linear combination of past values, where the weights are updated at every time step.²²
4. The wavelet least mean squares (WLMS) prediction filter decomposes first the signal into signal bands and then applies the least mean squares algorithm to each band separately. The results are then summed up.²³
5. The interacting multiple model (IMM) prediction filter combines two Kalman filters, the first one with a constant-velocity model and the second one with a constant-acceleration model.²⁴
6. The multistep linear (MULIN) prediction filter predicts using a procedure similar to a Taylor expansion around the current time step.²⁵
7. The support vector regression (SVR) prediction filter uses the SVR method, which allows for some non-penalizing deviations between the resultant function and the data²⁶. Here, the input data are the time indices and the output data consisting of the trace values corresponding to the time indices.
8. The goal of the kernel density estimation (KDE) prediction filter is to estimate a probability distribution using past values and, consequently, to predict the trace using the current values and the probability distribution.²⁷
9. The local circular model extended Kalman filter (LCMEKF) prediction filter uses a Kalman filter with a model that generates circular dynamics.²⁸
10. The recursive least squares (RLS) prediction filter works similar as the NLMS prediction filter but it differs in the way the weights are updated.²⁹
11. The extended recursive least squares (ERLS) prediction filter is essentially the same as the RLS prediction filter, but includes a parameter which allows to tune the prediction filter to the measurement noise.³⁰
12. The Takens (TAKENS) prediction filter is based on the Takens Reconstruction Theorem for nonlinear dynamics. Essentially, it compares the current values to past values and takes the closest past values to predict the respiratory motion.³⁰

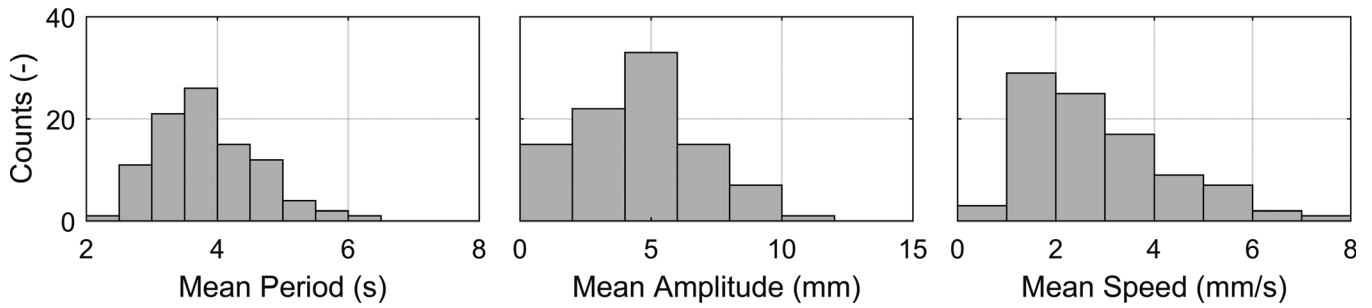


FIG. 2. Characteristics of respiratory motion traces. Left panel shows the distribution of the mean period of each trace. The center panel shows the distribution of the mean amplitude of each trace. The right panel shows the distribution of the mean speed of each trace.

13. The nonlinearly extended recursive least squares (NERLS) prediction filter combines the ERLS prediction filter with the TAKENS prediction filter. The ERLS predicts the trace, but simultaneously, its prediction errors are recorded. These are predicted by TAKENS and the predicted error is used to correct the ERLS prediction.³⁰
14. The Kalman filter constant acceleration (KFCA) prediction filter uses a Kalman filter with a constant acceleration model.³⁰
15. The Kalman filter constant velocity (KFCV) prediction filter uses a Kalman filter with a constant velocity model.³⁰
16. The relevance vector machine (RVM) prediction filter uses the RVM algorithm, which is based on a sparse Bayesian learning framework.³¹ The RVM prediction filter evaluates at every time step but trains in every tenth time step only.
17. The neural network (NN) prediction filter trains the NN using a sliding window of past values and uses the current values to predict the future values of the trace.³² The NN evaluates in every time step but trains in every tenth time step only to reduce computation time.
18. The linear filter (LF) prediction filter uses the past values to compute a linear regression which is then used with the current values to predict the future value.

The implementations of the prediction filters in MATLAB (The MathWorks Inc., Natick, MA, USA) are available in the supplementary material “Optimal Hyperparameters” and via the digital object identifier <https://doi.org/10.5905/ethz-1007-232>.

2.C. Prediction filter hyperparameter optimization

The prediction filters have various hyperparameters, which influence their performance to predict the respiratory motion. The performance of a prediction filter for a given trace was defined as follows: First, the root mean square (RMS) of the differences between the original signal x_{orig} and the prediction filter output signal x_{pred} was computed. Second, the RMS of the differences between the original signal x_{orig} and

the delayed signal x_{delayed} (delayed output of sensor) was computed. Finally, the ratio of these two values was assessed, which is denoted the normalized root mean square error e_{nRMS} . The e_{nRMS} was chosen because it is dimensionless and independent of the amplitude of motion traces. It directly indicates how the prediction filter’s performance compares to not using a prediction filter. An e_{nRMS} equal to 1 corresponds to no improvement, while an e_{nRMS} equal to 0 indicates perfect prediction. An e_{nRMS} above 1 indicates that using the prediction filter is actually worse than not using it. Because most prediction filters require a certain time to fully initialize, only the time after i_{cutoff} was taken into account for computing e_{nRMS} . Here, the first 20 s of each trace were ignored, corresponding to a $i_{\text{cutoff}} = 500$ steps. The variable N represents the length of the trace.

$$e_{\text{nRMS}} = \frac{e_{\text{RMS}}(x_{\text{orig}}, x_{\text{pred}})}{e_{\text{RMS}}(x_{\text{orig}}, x_{\text{delayed}})} = \frac{\sqrt{\frac{1}{N} \sum_{i > i_{\text{cutoff}}} (x_{\text{orig}}(i) - x_{\text{pred}}(i))^2}}{\sqrt{\frac{1}{N} \sum_{i > i_{\text{cutoff}}} (x_{\text{orig}}(i) - x_{\text{delayed}}(i))^2}}$$

We applied Bayesian optimization for the hyperparameter selection, implemented via the “bayesopt” function in the MATLAB Statistics and Machine Learning Toolbox (The MathWorks Inc., Natick, MA, USA), because it can handle continuous and discrete variables. The hyperparameters of each prediction filter were optimized with a subset of ten traces of the respiratory motion dataset introduced above. This subset was selected randomly exactly once for the entire study and substantially smaller than the full set to reduce the high computational effort required by the optimization algorithm and some by some prediction filters. The objective of the optimization algorithm was to minimize the median e_{nRMS} of the ten traces in the optimization subset. The optimization method used was the Bayesian Optimization³³ algorithm, which is widely used for the optimization of the hyperparameters of machine learning algorithms.

However, the optimal hyperparameters found were then evaluated on the validation subset, which consisted of the full set without the optimization subset. This entire procedure of optimizing the prediction filters’ hyperparameters and evaluating them was carried out for six scenarios. This number of scenarios results from the three different noise levels and the

two different time delays investigated in this study. The code for the optimization and the evaluation is available in the supplementary materials.

3. RESULTS

3.A. Performance of the prediction filter

The e_{nRMS} of all the prediction filters are shown as box plots in Fig. 3 (time delay = 160 ms corresponding to 4 time steps) and in Fig. 4 (time delay = 480 ms corresponding to 12 time steps). Additional results are presented in the supplementary materials. For both time delays, the first, second, and third panels show the performance for signals without noise, with an SNR ratio of 30 dB, and an SNR of 20 dB, respectively. Generally, smaller time delays and higher SNRs were beneficial for the performance of the prediction filters. Especially in the scenario of no noise, the WLMS and the LF prediction filter showed very small e_{nRMS} values for both time delay scenarios. But when noise was added, the results of these two prediction filters were closer to those of other prediction filters.

The prediction filters that reached the lowest median e_{nRMS} are shown in Table I for all scenarios of SNR and time delays. The best prediction filter for each scenario was determined as the one with the lowest median e_{nRMS} . The prediction filters with e_{nRMS} results comparable to the best one were determined by using non-inferiority tests³⁴ (one-sided Wilcoxon sign rank test with a tolerance value of 0.05). A prediction filter was non-inferior (or comparable) if its results were significantly below the best prediction filter's results plus a tolerance value of 0.05. This tolerance value is smaller than the interquartile ranges of the prediction filters' e_{nRMS} results in the noisy scenarios.

4. DISCUSSION

Tracking and compensating the motion of tumors can potentially reduce treatment margins and consequently reduce negative side effects. However, an effective tracking approach requires very small time delays from detecting the motion of the tumor to compensating the motion. The devices and approaches most commonly used in radiotherapy (motion sensor, signal processing, compensating motion) exhibit several time delays, which sum up to a large overall time delay. However, respiratory motion often has a regular pattern, which can be exploited by an algorithm that takes the past values of the motion signal to predict the signal ahead. Such algorithms are commonly denoted prediction filters.

Numerous prediction filters were proposed and investigated. In this paper, we compared 18 prediction filters on 93 respiratory motion traces with two time delays as well as three noise levels for each trace. The respiratory motion traces were standardized to a sampling frequency of 25 Hz and filtered with a cutoff frequency of 3 Hz. The hyperparameters of these prediction filters were optimized with a subset of ten traces. The prediction filters with the optimized

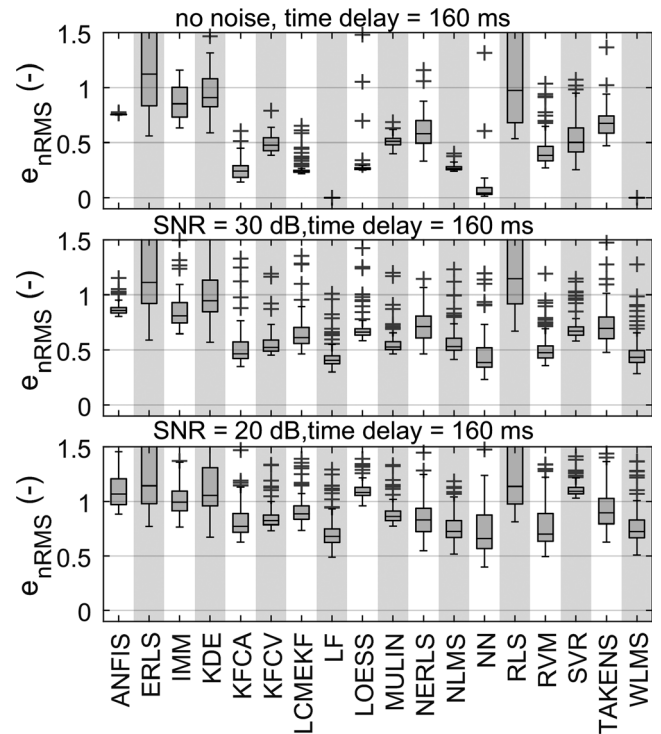


FIG. 3. Box plots of the performance (as measured by e_{nRMS}) of all optimized prediction filters for a time delay of 160 ms applied to the validation set of respiratory traces. The prediction filters are listed at the bottom and the abbreviations stated in the list of abbreviations. The first, second, and third panels show the performance for signals without noise, with a signal-to-noise (SNR) ratio of 30 dB and an SNR of 20 dB, respectively. The data points above an e_{nRMS} value of 1.5 are omitted. They are available in the supplementary material.

hyperparameters were then applied to the remaining 83 subset traces.

Generally, the e_{nRMS} value increases when the SNR decreases, which is in line with the results by Ernst et al.,¹⁶ who showed that smoothing the traces decreases the prediction errors. Prediction filters essentially extrapolate the signal, and extrapolation is very sensitive to noise. However, one would expect that increasing the time delay also increases the e_{nRMS} , but the results showed a mixed behavior among the prediction filters. For example, the e_{nRMS} value of the WLMS does not increase at all when the time delay is increased for smooth traces, but it does increase for noisy traces. The sensitivity of the WLMS to the time delay depends on the SNR. The KDE and the TAKENS show opposite behaviors of the e_{nRMS} , which decreased when the time delay was increased. This might be explained by the formula of the e_{nRMS} that corresponds to a ratio of two root mean square errors. A decrease of the e_{nRMS} cannot only be achieved with a decrease of the numerator, but also with an increase of the denominator. An increase of the time delay increases the value of the denominator. For the KDE and the TAKENS, the denominator's value increases more than the numerator's value as the time delay increases. This characteristic might make these prediction filters superior over the other prediction filters for longer time delays than those investigated here.

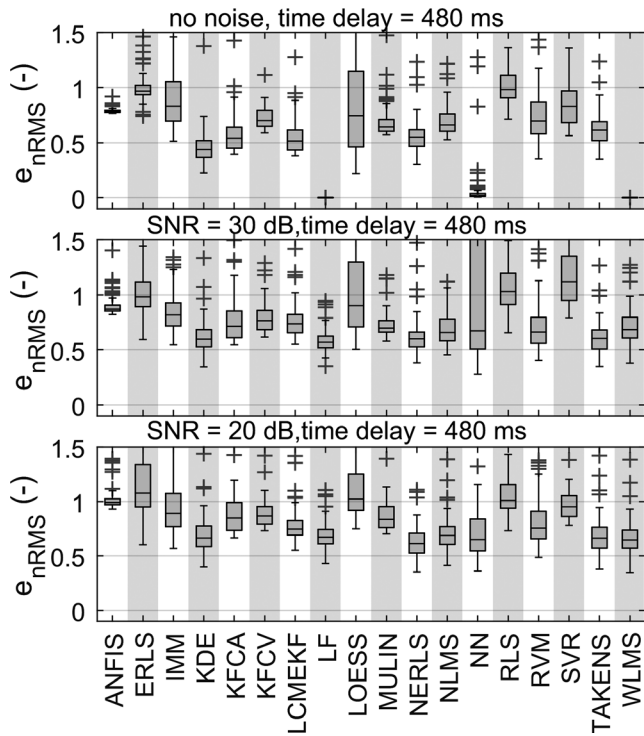


FIG. 4. Box plots of the performance (as measured by e_{nRMS}) of all optimized prediction filters for a time delay of 480 ms applied to the validation set of respiratory traces. The prediction filters are listed at the bottom and the abbreviations stated in the list of abbreviations. The first, second, and third panels show the performance for signals without noise, with a signal-to-noise (SNR) ratio of 30 dB, and an SNR of 20 dB, respectively. Data points above an e_{nRMS} value of 1.5 are omitted. They are available in the supplementary material.

TABLE I. Prediction filters with the lowest median as well as prediction filters whose results are significantly non-inferior (one-sided Wilcoxon sign rank test with a tolerance value of 0.05).

SNR	Time delay (ms)	Lowest median e_{nRMS} (-)	Prediction filter with lowest median e_{nRMS}	Non-inferior prediction filters
No noise	160	<0.01	WLMS	LF
30 dB	160	0.38	NN	LF, WLMS
20 dB	160	0.66	NN	LF
No noise	480	<0.01	WLMS	LF, NN
30 dB	480	0.56	LF	KDE, NERLS, TAKENS
20 dB	480	0.61	NERLS	-

The variance of the e_{nRMS} values in Figs. 3 and 4 for each prediction filter shows that the performance of the prediction filters is patient specific. In the scenario of smooth traces, the variance differs substantially among the prediction filters, whereas for noisy traces, the variances are more similar and increased for all prediction filters. Furthermore, the performance of a prediction filter depends more on the specific respiratory motion trace as the SNR decreases. The distributions of the e_{nRMS} values also show a large number of outliers,

which may indicate that there are always some traces, but not necessarily identical ones that are difficult for a given prediction filter.

There are several prediction filters, which showed remarkably low e_{nRMS} values in the scenario of smooth traces, especially the WLMS and the LF. Interestingly, these are both linear techniques. This observation indicates that linear methods are sufficient to describe respiratory motion and that non-linear approaches are not necessarily needed.

The large influence of the SNR points toward several possibilities to improve the performance of motion compensation. The sensor that measures the respiratory motion could be designed such as to achieve minimal noise. Another approach could be to include low-pass filters, which decrease the noise of the measured signal. However, a low-pass filter further increases the time delay, which would have to be compensated by the prediction filter in addition to the actual time delay.

The variation of the performance of a prediction filter over the various traces may indicate that for a given trace, an optimal prediction filter can be selected by considering the trace's characteristic features, which was investigated by Ernst et al.¹⁷ However, selecting a prediction algorithm for a given respiration pattern could be done with the algorithms themselves, because the prediction filters implemented here require only a few milliseconds for a time step to predict (see supplementary materials). Therefore, a patient's respiratory pattern could be recorded in advance and then the prediction filters could be tested on the respective recording. A 1-min recording tested with the LF would take roughly 2 s of computation time on a personal computer. Therefore, the prediction filters' results themselves can be used to select the best prediction filter for the patient.

Ernst and Schweikard¹² compared various prediction filters: The NLMS, RLS, WLMS, and MULIN prediction filters were investigated for a real respiratory motion signal including noise and a time delay of 150 ms. Their NLMS achieved an e_{nRMS} value of 0.95, which is in the range of the e_{nRMS} value for a time delay of 160 ms and an SNR of 20 dB shown in our study. The WLMS and the MULIN prediction filters of our study achieved lower median e_{nRMS} values (WLMS: 0.41, MULIN: 0.51 at 30 dB and 160 ms) than their WLMS and MULIN with 0.63 and 0.64, respectively. Our results are consistent with their results, but our results cover a larger variety of respiratory motion traces and prediction filters.

In more recent work by Ernst et al.,¹³ the MULIN, nLMS, RLS, wLMS, EKF, and SVRpred were compared to each other. Their results for a time delay of 154 ms and our results for an SNR of 30 dB and a time delay of 160 ms are similar, except for the RLS. It reached a lower e_{nRMS} than the RLS results of our study. Their MULIN has a mean e_{nRMS} of 0.61, while our results showed a median of 0.53. However, the MULIN's variance is asymmetric toward higher values, therefore, the mean is expected to be higher than the median. Their wLMS showed the lowest e_{nRMS} of 0.56, while our WLMS with an e_{nRMS} of 0.43 is comparable to the NN, see

Table I. The differences between their and our results may arise from the different time delays and noise conditions.

The KDE, LR (similar to LF), and NN prediction filters were investigated by Krauss et al.¹⁴ for time delays of 200, 400, and 600 ms, with resulting e_{nRMS} of 0.49, 0.54, 0.63, respectively. Their KDE prediction filter achieved substantially smaller e_{nRMS} values for time delays of 200 ms than the KDE achieved for 160 ms of our study with median e_{nRMS} close to 1.00. The difference decreases with longer time delays of above 400 ms, because at 480 ms, the KDE achieves median e_{nRMS} around 0.5. Their LR and NN prediction filters tend to reach lower e_{nRMS} values than their KDE, which also holds for our NN and LF prediction filters. Both in their work and our study, the LR has a median e_{nRMS} value around 0.5 for a time delay of 400 ms (their work) or our LF at 480 ms and 30 dB (our study). In this case, our results are mostly consistent except for the KDE for short time delays. This discrepancy may result from differences in the hyperparameter selection of the KDE prediction filter. The optimization algorithm used here does not guarantee that the global optimum was reached and can only search in a bounded range of values for the hyperparameters. Therefore, some of the prediction filters' performance results might not actually show their full potential. This could be mitigated by repeating the optimization with random starting hyperparameters and checking whether the optimal hyperparameters reach a bound.

An important performance criterion of a prediction filter is its behavior during breathing irregularities, which can result, that is, in baseline shifts, amplitude changes, or frequency changes. Examples of the behavior of all prediction filters for sudden baseline shifts and frequency changes are shown in the supplementary material "Supplement_Additional_Results" in the section "Breathing irregularities." Interestingly, the WLMS manages these irregularities with almost imperceptible errors in the no noise scenario. But it can also be observed that noisy traces decrease the performance of the WLMS during breathing irregularities. Therefore, smoothing the traces may improve the handling of irregularities.

The respiratory traces were standardized by Fourier transform and by removing all frequency components above 3 Hz. The resulting traces may not entirely correspond to the actual respiratory motion as it might contain frequencies above 3 Hz that are not present due to measurement noise. The magnitude of the added noise was defined by the SNR approach, where signals with smaller amplitudes have a smaller noise magnitude. This approach does not cover cases, in which the traces have different motion amplitudes but identical noise magnitudes. Furthermore, sensors also vary in their sampling frequencies. Here, only a sampling frequency of 25 Hz was considered.

5. CONCLUSIONS

For the dataset employed and using our implementations, we found that the best prediction filter was the linear filter. Reducing the measurement noise generally improves the performance of the prediction filters investigated in this study,

even during breathing irregularities. This finding highlights the relevance of future research on prediction filters toward increasing robustness against measurement noise and on motion measurement devices with minimal noise.

ACKNOWLEDGMENTS

We thank Samuel Dobmann and Yannick Berdou for their work on researching literature about prediction filters for human respiration prediction.

FUNDING

This work was supported by the Swiss National Science Foundation (SNSF) through "Development of prediction models for liver, lung, and breast tumors and implementation and verification of prediction filters for advanced couch tracking in a clinical environment," Grant No. CR32I3_153491.

CONFLICT OF INTEREST

The authors have no conflict to disclose.

^{a)}Author to whom correspondence should be addressed. Electronic mail: marischm@ethz.ch; Telephone: +41 44 632 24 47.

REFERENCES

- Dieterich S, Suh Y. Tumor motion ranges due to respiration and respiratory motion characteristics. In: Urschel HC, Kresl JJ, Luketich JD, Papiez L, Timmerman RD, Schulz RA, eds. *Treat. Tumors that Move with Respir.* Berlin Heidelberg: Springer; 2007:3–13.
- Rohlfing T, Maurer CR, O'Dell WG, Zhong J. Modeling liver motion and deformation during the respiratory cycle using intensity-based nonrigid registration of gated MR images. *Med Phys.* 2004;31:427–432.
- Van Den Begin R, Engels B, Gevaert T, et al. Impact of inadequate respiratory motion management in SBRT for oligometastatic colorectal cancer. *Radiother Oncol.* 2014;113:235–239.
- Keall PJ, Mageras GS, Balter JM, et al. The management of respiratory motion in radiation oncology report of AAPM Task Group 76a). *Med Phys.* 2006;33:3874–3900.
- Hoogeman M, Prévost JB, Nuytens J, Pöll J, Levendag P, Heijmen B. Clinical accuracy of the respiratory tumor tracking system of the cyberknife: assessment by analysis of log files. *Int J Radiat Oncol.* 2009;74:297–303.
- Depuydt T, Verellen D, Haas O, et al. Geometric accuracy of a novel gimbal based radiation therapy tumor tracking system. *Radiother Oncol.* 2011;98:365–372.
- Booth JT, Caillet V, Hardcastle N, et al. The first patient treatment of electromagnetic-guided real time adaptive radiotherapy using MLC tracking for lung SABR. *Radiother Oncol.* 2016;121:19–25.
- Lang S, Zeimet J, Ochsner G, Schmid Daners M, Riesterer O, Klöck S. Development and evaluation of a prototype tracking system using the treatment couch. *Med Phys.* 2014;41:21720.
- Tatinati S, Nazarpour K, Tech Ang W, Veluvolu KC. Ensemble framework based real-time respiratory motion prediction for adaptive radiotherapy applications. *Med Eng Phys.* 2016;38:749–757.
- Hong S-M, Bukhari W. Real-time prediction of respiratory motion using a cascade structure of an extended Kalman filter and support vector regression. *Phys Med Biol.* 2014;59:3555–3573.
- Choi S, Chang Y, Kim N, Park SH, Song SY, Kang HS. Performance enhancement of respiratory tumor motion prediction using adaptive

- support vector regression: comparison with adaptive neural network method. *Int J Imaging Syst Technol.* 2014;24:8–15.
12. Ernst F, Schweikard A. A survey of algorithms for respiratory motion prediction in robotic radiosurgery. *GI Jahrestagung.* 2009;1035–1043.
 13. Ernst F, Dürichen R, Schlaefer A, Schweikard A. Evaluating and comparing algorithms for respiratory motion prediction. *Phys Med Biol.* 2013;58:3911–3929.
 14. Krauss A, Nill S, Oelfke U. The comparative performance of four respiratory motion predictors for real-time tumour tracking. *Phys Med Biol.* 2011;56:5303–5317.
 15. Murphy MJ, Dieterich S. Comparative performance of linear and nonlinear neural networks to predict irregular breathing. *Phys Med Biol.* 2006;51:5903–5914.
 16. Ernst F, Schlaefer A, Schweikard A. Smoothing of respiratory motion traces for motion-compensated radiotherapy. *Med Phys.* 2010;37:282–294.
 17. Ernst F, Schlaefer A, Schweikard A. Predicting the outcome of respiratory motion prediction. *Med Phys.* 2011;38:417051118–14794497.
 18. Ernst F. *Compensating for Quasi-Periodic Motion in Robotic Radio-surgery.* Berlin: Springer Science & Business Media; 2012.
 19. Cho B, Poulsen PR, Sloutsky A, Sawant A, Keall PJ. First demonstration of combined kV/MV image-guided real-time dynamic multileaf-collimator target tracking. *Int J Radiat Oncol Biol Phys.* 2009;74:859–867.
 20. Kakar M, Aarup LR, Olsen DR. Respiratory motion prediction by using the adaptive neuro fuzzy inference system (ANFIS). *Phys Med Biol.* 2005;50:4721–4728.
 21. Ruan D, Fessler JA, Balter JM. Real-time prediction of respiratory motion based on local regression methods. *Phys Med Biol.* 2007;52:7137–7152.
 22. Ernst F, Schlaefer A, Dieterich S, Schweikard A. A fast lane approach to LMS prediction of respiratory motion signals. *Biomed Signal Process Control.* 2008;3:291–299.
 23. Ernst F, Schlaefer A, Schweikard A. Prediction of respiratory motion with wavelet-based multiscale autoregression. In: Ayache N, Ourselin S, Maeder A, eds. *International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007.* Berlin, Heidelberg: Springer; 2007:668–675.
 24. Putra D, Haas O, Mills JA, Burnham KJ. A multiple model approach to respiratory motion prediction for real-time IGRT. *Phys Med Biol.* 2008;53:1651–1663.
 25. Ernst F, Schweikard A. Predicting respiratory motion signals for image-guided radiotherapy using multi-step linear methods (MULIN). *Int J Comput Assist Radiol Surg.* 2008;3:85–90.
 26. Ernst F, Schweikard A. Forecasting respiratory motion with accurate online support vector regression (SVRpred). *Int J Comput Assist Radiol Surg.* 2009;4:439–447.
 27. Ruan D. Kernel density estimation-based real-time prediction for respiratory motion. *Phys Med Biol.* 2010;55:1311–1326.
 28. Hong S-M, Jung B-H, Ruan D. Real-time prediction of respiratory motion based on a local dynamic model in an augmented space. *Phys Med Biol.* 2011;56:1775–1789.
 29. Ernst F, Schweikard A. Prediction of respiratory motion using a modified Recursive Least Squares algorithm. in CURAC; 2008.
 30. Herrmann C. *Robotic Motion Compensation for Applications in Radiation Oncology.* Würzburg: Uni Würzburg; 2013.
 31. Dürichen R, Wissel T, Ernst F, Schweikard A. Respiratory Motion Compensation with Relevance Vector Machines Robert. In International Conference on Medical Image Computing and Computer-Assisted Intervention 2013. Berlin, Heidelberg: Springer; 2013:108–115.
 32. Teo PT, Bruce N, Pistorius S. Application and parametric studies of a sliding window neural network for respiratory motion predictions of lung cancer patients. *IFMBE Proc.* 2015;595–598.
 33. Shahriari B, Swersky K, Wang Z, Adams RP, de Freitas N. Taking the human out of the loop: a review of Bayesian optimization. *Proc IEEE.* 2016;104:148–175.
 34. Walker E, Nowacki AS. Understanding equivalence and noninferiority testing. *J Gen Intern Med.* 2011;26:192–196.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Data S1: Further results.

Data S2: Detailed hyperparameter values.

Data S3: Source Code.