Title

Parameter Selection for Convex Optimization Time Calibration for a 2-Panel PET System.

Authors

Romanchek, Gregory

Abbaszadeh, Shiva

Publication Date

2023-02-01

DOI

10.1117/12.2654423

Peer reviewed

HHS Public Access

Author manuscript

Proc SPIE Int Soc Opt Eng. Author manuscript; available in PMC 2024 November 02.

Published in final edited form as:

Proc SPIE Int Soc Opt Eng. 2023 February ; 12463: . doi:10.1117/12.2654423.

Parameter Selection for Convex Optimization Time Calibration for a 2-Panel PET System

Gregory Romanchek[a,b], **Shiva Abbaszadeh**[b]

[a]Dept. of Nuclear, Plasma, and Radiological Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61891, USA;

[b]Dept. of Electrical and Computer Engineering, University of California Santa Cruz, Santa Cruz, CA 95064, USA

Abstract

Coincidence timing calibration is fundamental to PET imaging. The electronics, cable lengths, and detector physics such as charge drift and depth dependence add to the measured time differences in coincidence sorting – increasing random rate, decreasing true rate, and degrading system performance. This work investigates the parameter selection for convex optimization (Ordinary Least Squares) for timing calibration. We test the correlation between commonly selected parameters and the experimentally measured coincidence time difference. Additionally, we test 127 nested models of a parameterized regression equation to identify the those which optimize MSE, BIC, and FWHM, respectively. In each of these models, the FWHM performance improved ~53%, though the value shifted from ~ 304 to 160 ns – far from ~ 10 ns FWHM CZT can achieve. The results point to the lack of a necessary parameter, such as trigger threshold level or temperature, or data which is too variable for the OLS optimization.

Keywords

PET; time resolution; time calibration; convex optimization; parameter selection

1. INTRODUCTION

In Positron Emission Tomography (PET) imaging, adequate timing resolution (towards tens of picoseconds for time-of-flight (ToF) systems and within few nanoseconds in non-ToF systems) is necessary for coincidence assignment, random rejection, and ToF imaging. However, system intrinsic and physical sources of temporal delay degrade the timing resolution of modern systems. The experimentally measured time difference of a coincidence pair ΔT_{exp} is thus the combination of factors beyond the true time-of-flight difference from point of annihilation to detector element. Such factors can include the time delay from charge drift ΔT_{drift}, channel crosstalk causing the signal trigger voltage threshold to fluctuate $\Delta T_{crosstalk}$, the finite speed of electromagnetic waves in hardware (system) ΔT_{delay}, the actual time of flight from point of annihilation to detection (often ignored due to negligible magnitude), and more. The compounding of these delays alters the true time difference ΔT_{true}:

$$\Delta T_{exp} = \Delta T_{true} + \Delta T_{drift} + \Delta T_{delay} + \Delta T_{crosstalk} \quad (1)$$

In essence, timing calibration is the removal of these sources of delay to recapture the true coincidence ΔT. As such delays differ stochastically for each coincidence detection and channel pair, the true (read, unknowable) value are replaced with estimates:

$$\widehat{\Delta T}_{true} = \Delta T_{exp} - \widehat{\Delta T}_{drift} - \widehat{\Delta T}_{delay} - \widehat{\Delta T}_{crosstalk} \quad (2)$$

$$\widehat{\Delta T}_{true} = \Delta T_{exp} - \widehat{\Delta T}_{error} \quad (3)$$

While many methods exist to correct for delay effects, it has been shown such timing calibration can be formulated as a convex optimization problem utilizing any number of parameters [1]. Thus, the solution can be found with well-known optimization methodologies. For systems similar to the target UCSC 2-Panel PET Scanner, typical parameters often include the ADC values of the analog quadrature timing signals of the application specific integrated circuits (ASIC) (described below), their combinations, and interaction depth, such as through the deposited energy readout out from cathode-to-anode ratio C/A. A few examples of models are presented below:

$$\widehat{\Delta T}_{error} = a_0 U + a_1 V + a_2 U^2 + a_3 V^2 + a_4 UV + a_5 \quad (4)[1]$$

$$\widehat{\Delta T}_{error} = a_0 U_c + a_1 U_c^2 + a_2 V_c + a_3 V_c^2 + a_4 + a_5 U_a + a_6 U_a^2 + a_7 V_a + a_8 V_a^2 + a_{11} + a_9 \left(C/A\right) + a_{10} \left(C/A\right)^2 \quad (5)[2]$$

$$\widehat{\Delta T}_{error} = a_{0,i} U_i + a_{1,i} U_i^2 + a_{2,i} V_i + a_{4,i} (C/A)_i + a_{5,i} (C/A)_i^2 + a_{6-44,i} ASel_i - \left(a_{0,j} U_j + a_{1,j} U_j^2 + a_{2,j} V_j + a_{4,j} (C/A)_j + a_{5,j} (C/A)_j^2 + a_{6-44,j} ASel_j\right) \quad (6)[3]$$

where a_i are the coefficients to solve for, U and V are the quadrature timing signals, and ASel is a vector to select for the anode where the interaction occurred. Here, U and V are sinusoids (with V lagging by 90°) with a signal frequency of 490 kHz (though this can vary from system to system). These waves are sampled at the time of each event such that phase difference between their U+iV vectors can be computed. This angle, ψ, in conjunction with the UV-Frequency, f, is used to compute a fast timing signal, ΔT_{exp}, as shown in Eq. 7. The U and V signals from a single channel are plotted in Fig. 1.

$$\Delta T_{exp} = \left(\frac{\psi}{2\pi}\right)\left(\frac{1}{f}\right) \tag{7}$$

The diversity of construction of parameters in Eqs. 4–6 can be noted, especially as these were all formulated for low panel count, high resolution systems. This work strives to optimize the parameter selection and model construction for such systems. Failing to thoroughly investigate the model parameters can lead to slower optimization time, overfitting and poor generalizability, and overweighted parameter impact from collinearity between predictors. The inclusion of one additional parameter in the general formula increases the total number of parameters to solve for by the number of channels – which can number in the hundreds of thousands to millions – as each channel must be individually calibrated. Finally, complex models are less explainable than simple models – meaning abstraction back to the sources of delay in Eq. 1 may not be possible. In this work, we perform a parameter selection analysis on a series of models based up those collected from literature.

2. METHODS

2.1 System Design and Data

In this work, we are utilizing a PET system based on cadmium zinc telluride (CZT) with a RENA-3 ASIC. Data was obtained with a sub-system of this CZT-based design [4, 5]. The full scanner contains two panels of 5 × 30 stacked CZT detectors oriented in an edge-on configuration. Each of these detector modules contains two 40 × 40 × 5 mm^3 CZT crystals with cross-strip designs: 39 anode strips (100 μm width and 1 mm pitch) and 8 cathode strips (4900 μm width and 5 mm pitch). As such, each crystal has 47 output signals which are read out using two RENA-3 ASICs (by NOVA R&D Inc., Riverside, CA) which digitizes the signal, provides a coarse time stamp, and generates the phase-locked sine signals U and V. The sub-system used here contains 506 cathode-anode channel pairs across the two panels. The system design is highlighted in Fig. 2. The experimental data was collected by placing a Ge-68 point source at the midpoint of the scanner field of view. A visualization of the experimental ΔTs is shown in Fig. 3 and of the parameters of interest in Fig. 4.

2.2 Data Treatment

Mapping from ADC units to keV, U-V phase distortion correction and center point extraction, and anode signal amplitude deficit calibration were performed with the software toolkits described in [2]. Coincidences were sorted based on the coarse time stamp with a window of 1 μs. From these coincidences, the experimental time differences ΔT_{exp} for cathode-cathode events are computed via Eq. 7. In this system, we utilize cathodes as opposed to anodes for coincidence readout because they are wider (4900 μm for cathode and 100 μm for anode) and are triggered much faster during photon interactions within CZT crystals.

This full set data is subject to randoms and scatters, which introduce erroneous ΔT_{exp}s to our target set. Further, the parameters chosen (summarized below) vary in magnitude greatly and are not zero-mean – two aspects which degrade or otherwise unsubstantiate the regression optimization performed below. To treat the ΔT_{exp}, a simple outlier rejection is performed by fitting the ΔT_{exp} histogram with a Gaussian and discarding events where $\Delta T_{exp,i} > 3\sigma -$ Fig. 3 shows this data after such treatment. The collected regression parameters are z-score normalized to ensure similar value magnitude, allow for the solved regression coefficients to be more easily interpreted, and to ensure they are zero-mean. Additionally, channels with fewer counts than coefficients to solve for are discarded from the data set.

2.3 Model and Optimization

The full regression model selected to estimate $\widehat{\Delta T}_{error}$ is given by:

$$\widehat{\Delta T}_{error} = a_1 U_c + a_2 U_c^2 + a_3 U_c \cdot V_c + a_4 V_c + a_5 V_c^2 + a_6 \left(C/A\right) + a_7 \left(C/A\right)^2 + a_8 \tag{8}$$

We note “full model” here as the many tested models are nested versions of this one. In other words, each other model contains a subset of the parameters found in Eq. 8 with no new ones. Every combination of these parameters was tested (except always leaving a_8 as a constant) by dropping terms from this general form. Independent optimization was performed over these 127 nested models to test their performances.

The convex optimization process [1] is formulated as an Ordinary Least Squares (OLS) regression solving:

$$C_{opt} = \underset{C}{\operatorname{argmin}} \| \Delta T_{exp} - AC \|_2^2 \tag{9}$$

Here, $A \in \mathbb{R}^{n_c \times n_p}$ is the parameter matrix where n_c is the number of coincidence events in the data set and n_p is the number of channels times the number of parameters per channel. In a simplified case of a two-channel system and the model described in Eq. 8, for row 1, columns 1 – 8 would contain the parameter values for channel 1 coincidence 1 and columns 9 – 16 would contain the parameter values for channel 2 coincidence 1. Each channel has a full set of coefficients to solve for. In a system with more than 2 channels, the columns associated with the channels not participating in the event are set to 0. This way, the coefficient matrix $C \in \mathbb{R}^{n_p \times 1}$, and each parameter for each channel can be solved for at once. Namely, C_{opt} is directly solvable via the pseudo-inverse:

$$C_{opt} = \left(A^T \cdot A\right)^{-1} \cdot A^T \cdot \Delta T_{exp} \tag{10}$$

2.4 Training and Testing Sets

The training data set consisted of the channel pair with the greatest number of coincidence events between them. We are not solving for general coefficients here but rather which parameters to use in the general case. Limiting the test set to one channel pair allows for simpler analysis of the 16 total parameters rather than the hundreds involved in the full set. The testing set consisted of the total data of all channel pairs, ~35× larger than the training set.

2.5 Test Metrics

Performance of each model was gauged on three metrics: MSE (Mean Squared Error), BIC (Bayesian Information Criteria), and FWHM (Full Width Half Max). For each, smaller values indicate better performance. While MSE and BIC are standard statistics for model fitting, the FWHM is included as it is the performance metric of operational consequence. We are targeting a model which yields the best timing resolution as measured by the FWHM of the calibrated $\widehat{\Delta T}_{true}$. Additionally, the covariances of the parameters with each other and their partial variance with the target ΔT_{exp}. Finally, the coefficients, their variances across channels, and their standard errors are probed.

3. RESULTS

First, the covariance matrix in Table 2 also captures the Pearson correlation matrix as the parameters are z-score standardized. The diagonal, however, contains the variance of the corresponding parameter. The final column is the partial covariances between the parameters and the target ΔT_{exp}. These values are from the panel 1 value of the training data set. Mean centered on zero, the range of the U and V data is from ~ −0.8 to 0.8 while the C/A data is from ~ −1.5 to 1.5, meaning the variances and correlations are quite small among parameters. The largest correlation is between C/A and its transformation $(C/A)^2$ at 0.316. The partial correlations with ΔT are of a different scale as the target is not standardized. With this, we can observe strong correlations between U and C/A with the target and moderate to weak for the remainder.

Three separate models were found, one to optimize each of MSE, BIC, and FWHM for the training (one channel pair) data. The scores on the selected test metrics as well as the chosen variables are shown in Table 2. Additionally in Table 2, the performance of that model's parameter choice is given as applied to the testing set. The model selecting for MSE retained all parameters, which makes sense given the metric simply minimizes error. The BIC stat, which accounts for model complexity, threw out the U^2 parameter but kept the remainder. Physically, this finding may make sense as U^2 and V^2 can be thought of as surrogates for cos and sin – meaning the sum of their squares hold little information. The model which prioritized FWHM minimization kept only 5 of the 8 parameters, discarding U^2, V^2, and $U \cdot V$ – akin to removing 37.5% of the available parameter data.

Looking at the generalized performance of the model structures (not coefficients) on the testing set reveals a swap in performance for the BIC and FWHM models. Channel-to-channel variations exits, and so, for example, U may be of great importance to one set

of parameters while V hold that position for another set. In all cases, the FWHM of the experimental ΔTs was improved nearly two-fold from the 303.1 ns FWHM to ~ 158 ns. The models' application to the data can be seen in Fig. 5. This resolution is far from the ~10 ns FWHM CZT can achieve [6] – pointing to either a confounding variable or sub-optimal data quality. In [1], the uncalibrated to calibrated timing performance moved from 16.3 to 6.92 ns (a 57% decrease), on par our level of performance improvement.

Finally, for the testing set, we show the regression coefficients' mean and standard deviations in Table 3. Shown is also the mean standard error and its standard deviation for each parameter. The C/A ratio was assigned the largest coefficient on average while U^2 and V^2 had the largest standard errors on average. In all cases, however, the spread of the coefficients and standard errors are quite large compared to the magnitude of the coefficients themselves. This behavior again points to either poor input data quality, confounding variables, or both.

4. CONCLUSION

In this study, we optimized the timing calibration regression equation given the test metrics of MSE, BIC, and FWHM. Three parametric equation constructions were identified based on their performance of the three test statistics, and their performances were generalized to a larger testing set. Their generalizability did not hold, yielding poorer test statistic performance in all cases. The coefficients themselves varied greatly among the channels and their average standard errors were large compared to the coefficient magnitude except C/A. While the optimization improved FWHM from ~300 to ~160 ns (matching the percent performance improvement in literature), the obtained FWHM is far from that expected. Either a confounding variable exists, such as temperature or trigger threshold level, or the data at hand is too variable for such a simple optimization technique.

REFERENCES

[1]. Reynolds Paul D., Olcott Peter D., Pratx Guillem, Lau Frances WY, and Levin Craig S.. "Convex optimization of coincidence time resolution for a high-resolution PET system." IEEE transactions on medical imaging 30, no. 2 (2010): 391–400. [PubMed: 20876008]

[2]. Gu Yi. High-resolution small animal positron emission tomography system based on 3-d position-sensitive cadmium zinc telluride photon detectors. Stanford University, 2014.

[3]. Abbaszadeh Shiva, Gu Yi, Reynolds Paul D., and Levin Craig S.. "Characterization of a sub-assembly of 3D position sensitive cadmium zinc telluride detectors and electronics from a sub-millimeter resolution PET system." Physics in Medicine & Biology 61, no. 18 (2016): 6733. [PubMed: 27551981]

[4]. Wang Yuli, Herbst Ryan, and Abbaszadeh Shiva. "Development and Characterization of Modular Readout Design for Two-Panel Head-and-Neck Dedicated PET System Based on CZT Detectors." IEEE Transactions on Radiation and Plasma Medical Sciences 6, no. 5 (2021): 517–521. [PubMed: 37711549]

[5]. Li Mohan, Yockey Brett, and Abbaszadeh Shiva. "Design study of a dedicated head and neck cancer PET system." IEEE transactions on radiation and plasma medical sciences 4, no. 4 (2020): 489–497. [PubMed: 32632397]

[6]. Meng Ling J., and He Zhong. "Exploring the limiting timing resolution for large volume CZT detectors with waveform analysis." Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 550, no. 1–2 (2005): 435–445.

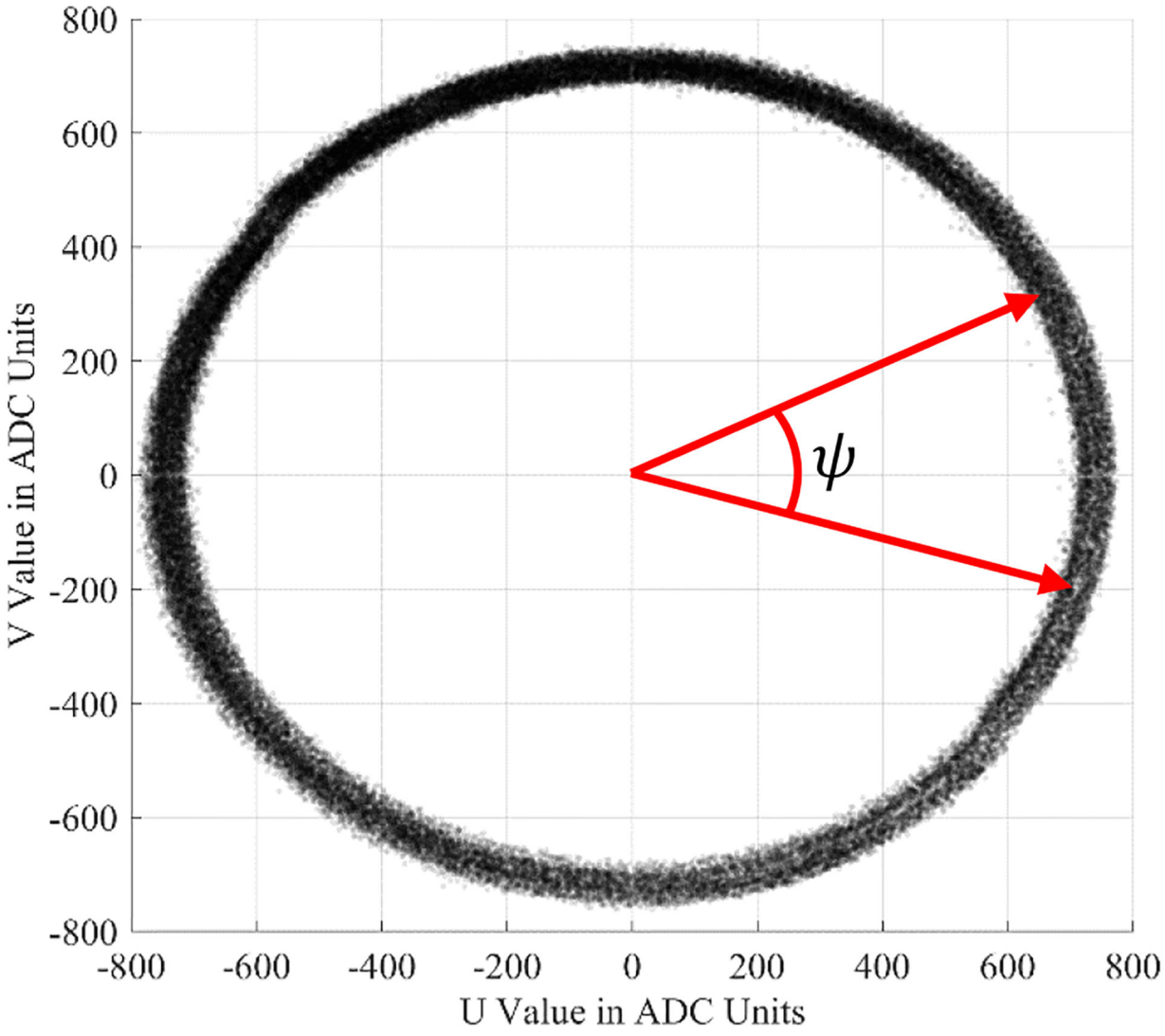

Figure 1.

V vs U values from one panel with an average radius of 729.17 and standard deviation (thickness of the circle perimeter) of 19.22. Each plotted dot captures the U and V values of a single detection event. If the red arrow indicate two events which were detected in coincidence, the angle between their U, V vectors is ψ.

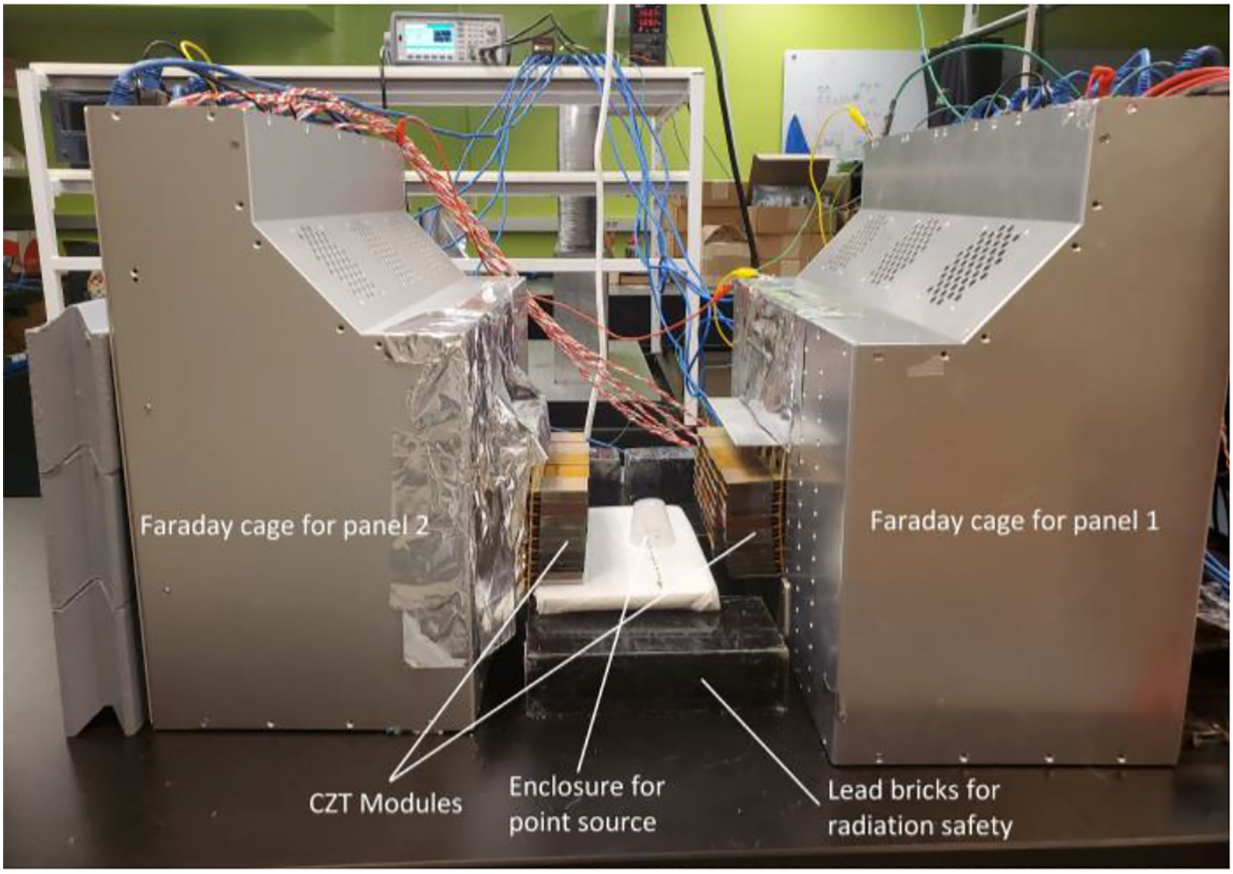

Figure 2.

We are utilizing a PET system based on CZT with a RENA-3 ASIC. Each of the CZT uses a cross-strip configuration consisting of 39 anode strips and 8 cathode strips. A description of the detector system and design is found in [3–5].

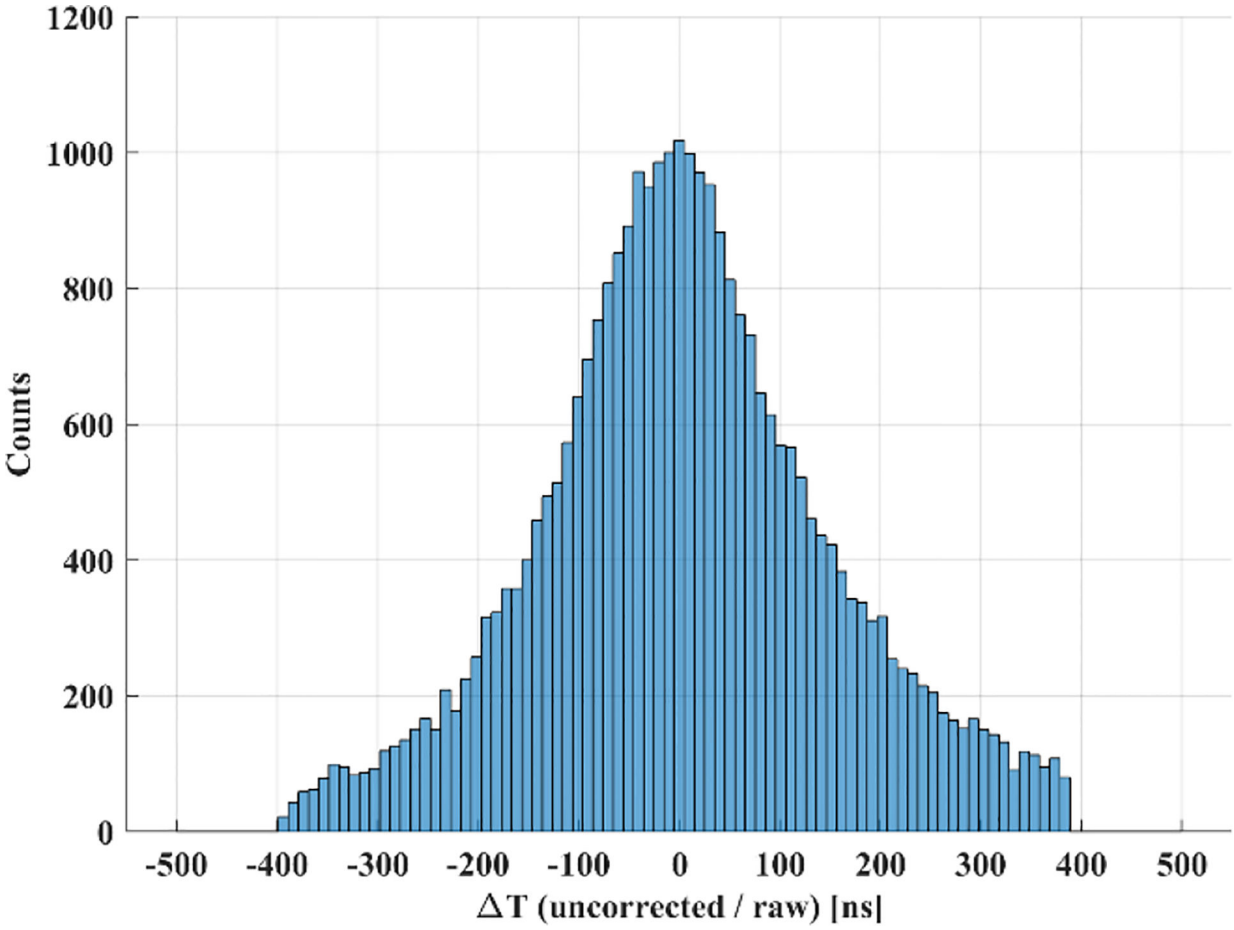

Figure 3.

Histogram of the uncalibrated ΔTs after coarse outlier removal. While centered near the ideal of 0 ns at −1.93 ns, the spread of the distribution (FWHM of 302.1 ns) is attributed to the delays discussed in Eq. 1 and is to be corrected.

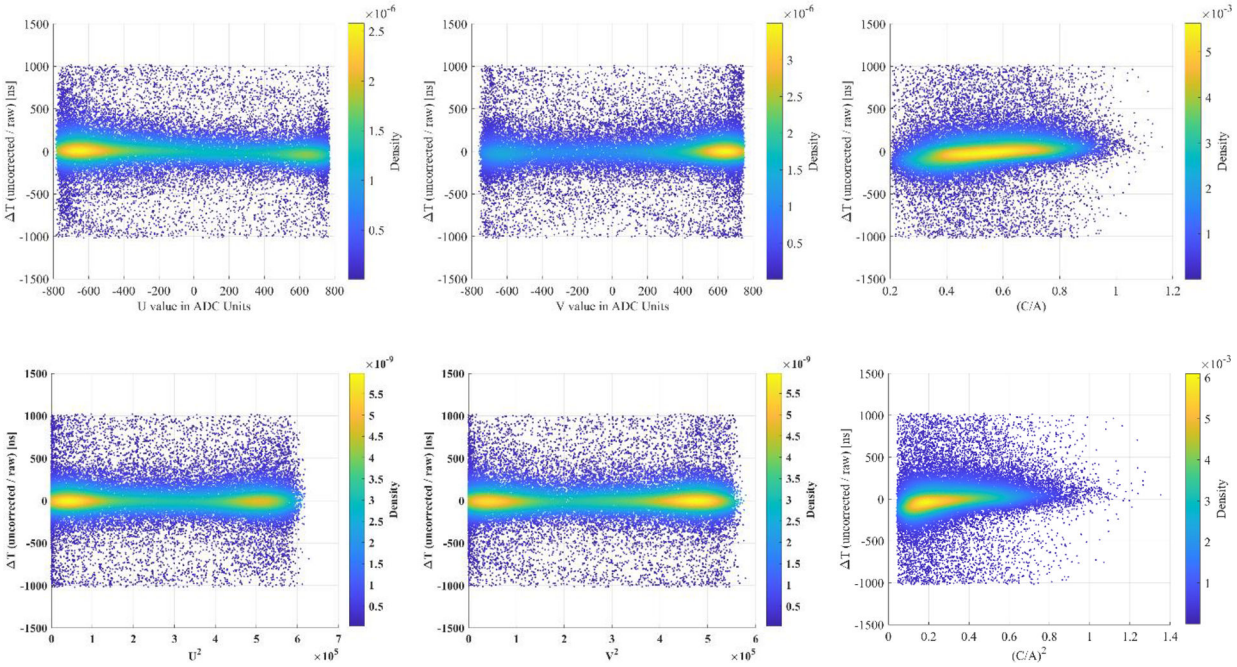

Figure 4.

Data histograms of the uncalibrated ΔTs vs the parameters U, V, C/A, U^2, V^2, and $(C/A)^2$ from panel 1. A density map is overlayed atop the scatter to better visualize the structure of the data. These represent the input parameters to the optimization problem at hand.

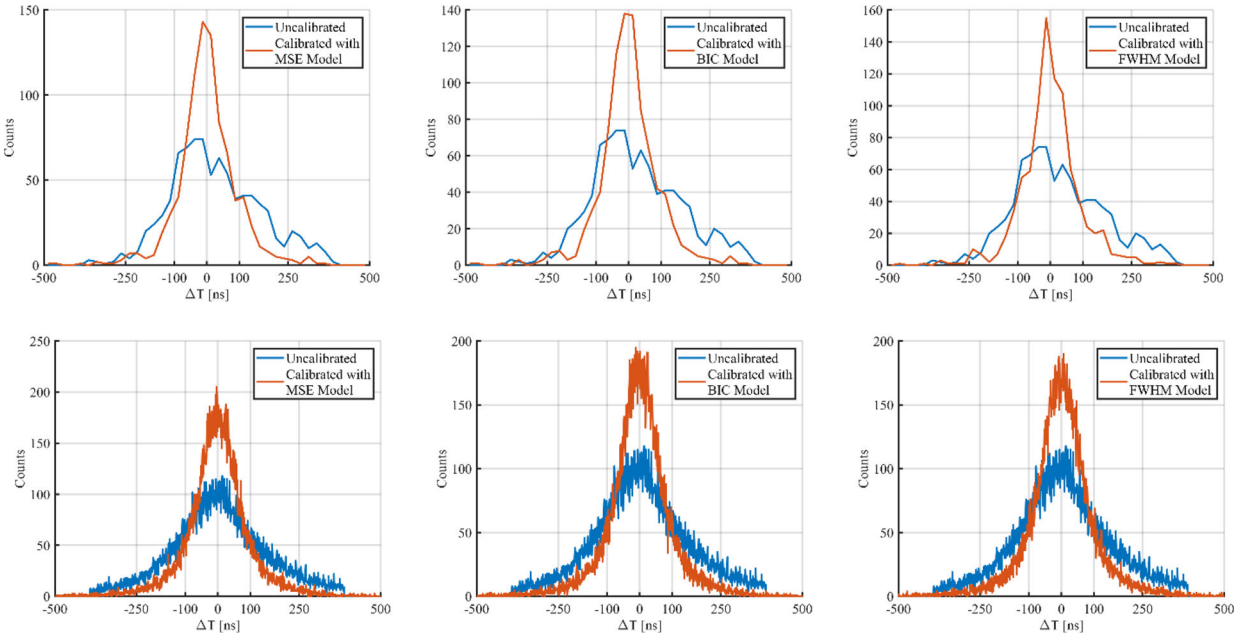

Figure 5.

Model performances of the three formulations which optimized MSE, BIC, and FWHM. The top row is the training set performance, and the bottom row is the testing set. The calibrated ΔT (orange) are plotted against the uncalibrated ΔT (blue).

Table 1.

Covariance between treated parameters and partial variance with ΔT. Between the parameters, small magnitudes are preferred meaning the parameters supply different information structure. Against ΔT, larger magnitudes are preferred. The diagonal is the variance of the parameter.

	U	**U^2**	**U · V**	**V**	**V^2**	**C/A**	**$(C/A)^2$**	**ΔT**
U	0.216	−0.003	0.025	−0.154	−0.001	−0.006	−0.003	−5.558
U^2		0.193	−0.169	−0.041	0.191	0.004	0.007	1.489
U · V			0.184	0.020	−0.171	0.001	0.000	−0.346
V				0.210	−0.042	0.009	−0.006	0.450
V^2					0.194	0.003	0.006	1.374
C/A						0.316	0.316	−2.363
$(C/A)^2$							0.349	−0.738

Table 2.

Optimized model performances. Bolded text highlights the minimum value (optimum) of each test criteria (each column).

		Training Set Metrics			Testing Set Metrics		
Optimized	**Parameters Selected**	**MSE**	**BIC**	**FWHM**	**MSE**	**BIC**	**FWHM**
MSE	Full Model	**8357.74**	8028.56	144.08	**9825.06**	288405.1	156.41
BIC	All but U^2	8360.30	**8015.28**	140.90	9860.02	287718.6	**155.95**
FWHM	U, V, C/A, $(C/A)^2$, 1	8750.20	8028.15	**138.35**	10373.79	**287685.9**	166.95

Table 3.

The standard errors are quite large, as are the variances in the parameter coefficients. The C/A ratio is quite prominent.

		Coefficient		Standard Error	
	For	**Mean**	**StDv**	**Mean**	**StDv**
$\mathbf{a}_1$	U	−14.17	21.46	8.78	11.30
$\mathbf{a}_2$	U^2	17.38	122.99	86.61	67.42
$\mathbf{a}_3$	$U \cdot V$	−5.18	22.29	7.72	7.96
$\mathbf{a}_4$	V	−38.12	38.84	9.22	14.91
$\mathbf{a}_5$	V^2	16.89	133.96	87.35	69.35
$\mathbf{a}_6$	C/A	127.93	78.94	41.77	34.75
$\mathbf{a}_7$	$(C/A)^2$	−66.24	71.97	41.70	35.19