NOAA Technical Report OTES 1

Error Analysis of Pulse Location Estimates for Simulated Bathymetric Lidar Returns

Rockville, Md.
July 1981

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
Ocean Technology and Engineering Services

NOAA Technical Report OTES 1

Error Analysis of Pulse Location Estimates for Simulated Bathymetric Lidar Returns

Gary C. Guenther
Engineering Development Office

Robert W. L. Thomas
EG&G/Washington Analytical Services Center, Inc.

Rockville, Md.
July 1981

U.S. DEPARTMENT OF COMMERCE
Malcolm Baldrige, Secretary

National Oceanic and Atmospheric Administration
John V. Byrne, Administrator

Ocean Technology and Engineering Services
M. E. Ringenbach, Director

Mention of a commercial company or product does not constitute an endorsement by NOAA/OTES. Use for publicity or advertising purposes of information from this publication concerning proprietary products or the tests of such products is not authorized.

CONTENTS

FIGURES

ABSTRACT

A Poisson count simulator has been exercised to generate precision and offset results for the estimated temporal location of quantum limited laser pulses. Pulse sizes, shapes, and charge integration times are varied over appropriate ranges. A number of location estimators including variations on peak, centroid, and threshold detectors are examined. Comparisons with experimental results are presented. Hardware and software design parameters for an airborne lidar hydrography system are discussed.

INTRODUCTION

The utilization of an airborne, scanning, pulsed lidar system for acquiring bathymetric data from coastal waters, bays, and lakes has been extensively studied in the past decade (Hickman and Hogg 1969, Cunningham 1972, Hickman et al. 1974, Abbot and Penny 1975, Kim et al. 1975, Goodman 1975, Kim and Ryan 1975, Witt et al. 1976, Goodman 1976, Goodman 1978, Guenther and Goodman 1978, Abbot et al. 1978, Thomas and Guenther 1979, Enabnit and Nield 1980, Hoge et al. 1980, Guenther and Thomas 1980). Experimental programs with second-generation hardware are currently underway in Canada and Australia as well as in the United States (Guenther and Goodman 1979, Ryan and O'Neil 1980, Van Norden 1980, Guenther and Thomas 1981). This technique is destined for implementation and integration into the spectrum of existing bathymetric systems.

The ultimate operational feasibility of airborne lidar hydrography depends on a number of physical, economic, and logistical factors. The results of several cost studies (i.e., Enabnit and Goodman et al. 1978) have indicated that the augmentation of existing sonar capabilities in "shallow" water with a single airborne laser system should result in substantial savings in operational costs, and the acquisition costs would be recovered in less than two years. Logistical requirements such as navigation, positioning, water clarity forecasting, tide control, data reduction, etc., are being actively studied and are not expected to provide any major barriers.

The two basic physical constraints on a system utilizing this technology are depth measurement "accuracy" and maximum "penetration" depth. The accuracy of depth measurements encompasses several concepts such as precision, repeatability, and comparability with independent standards. Penetration implies not just the capability to detect but also to accurately locate a return signal which has traversed the water column, been reflected from the bottom, again traversed the water column, and arrived as a distant receiver. Useful penetration is limited, therefore, not by detection, but by the accuracy requirement. Accuracy is the paramount concept upon which the design of such a system must be predicated.

The basic accuracy standard in the United States for hydrography in "shallow" water (less than 20 m) is ±30 cm. This is the total permissible error accumulated from all sources including random error, bias, tide control, positioning, waves, etc. It is estimated that a maximum of roughly half this error budget, ±15 cm, should be allocated for "precision," i.e., lack of short-term (pulse to pulse) repeatability. This is equivalent to ±1.3 nanoseconds (ns) of round-trip time in water. In this domain (gigahertz frequency) it may be perceived that extremely stringent requirements are incumbent on the receiving electronics system and the processing algorithms - particularly in view of the fact that a practical laser pulse is expected to be emitted with a "half width" (FWHM) of no less than about 5.0 ns and will be further stretched as it propagates through the water. Complementary approaches have been invoked in a comprehensive study of the errors associated with airborne lidar bathymetry over a range of depth and water clarity conditions: the field test of a research instrument - NASA's Airborne Oceanographic Lidar system (Guenther and Goodman 1978, Guenther 1978), a derivation of pulse stretching and return power based on scattering and diffusion theory (Guenther and Thomas 1979), a Monte Carlo propagation simulation (Guenther and Thomas 1981), and the computer simulation described in this report to predict, verify, and extrapolate experimental results for "random" error.

The detailed design of a prototype airborne lidar bathymetry system (consisting of a laser transmitter, scanner and receiver optics, data acquisition electronics, data recording, analysis, and control computer, and integrated attitude, navigation, and positioning electronics) is extremely complex and involves a number of difficult hardware tradeoffs at a state-of-the-art level. It is absolutely vital that no hardware configuration be selected before a careful analysis is performed to determine the impact of that design on the inherent accuracy potential of the instrument.

The purpose of this study is to provide a standard upon which such judgements can be based. The simulations herein described have been conducted with algorithms and parameter ranges consistent with advanced state-of-the-art laser and electronic hardware in such a way that the effects on the accuracy of algorithm or parameter trade-offs can be explicitly resolved. In this way, development programs for unavailable components can be prioritized not only by their cost but by their cost-effectiveness. Furthermore, pulse location techniques, whether implemented in hardware or software, analog or digital, may be selected to provide the optimum balance between precision and bias errors over a typical range of operating conditions. This will permit the selection of the best of possible alternatives within the constraints of resources, technology, and operational restrictions.

FORMAL STATISTICAL DESCRIPTION

The concept to be simulated consists of an input pulse waveform sampled by a series of discrete, temporally adjoined quantum accumulation intervals which, in effect, integrate charge from adjacent time slices across the superimposed pulse. In practice, this could be the results of sampling a pulsed lidar signal from the output of the photocathode of a photomultiplier tube with an a/d converter. The simulation is performed at the output of the photocathode since this location represents the lowest signal level and thus dictates the limiting signal-to-noise ratio for the system. The resultant output is a series of cardinal numerals representing the time sampled amplitude variation of the input pulse. These integer values are subject to Poisson statistics; that is to say, for any fixed mean input signal or "rate" in a time slice, the distribution of output values from that time "bin", over a number of pulses, will be Poisson. When large mean rates are involved, Gaussian statistics provide a good approximation to Poisson statistics.

The purpose of this section is to generate a mathematical description of the applicable Poisson statistics and to introduce several general categories of pulse location tests for interpreting the data. We also demonstrate how the probability distributions of the simpler test statistics can be derived formally.

Suppose that the expected count in the k^{th} time bin is λ_k. This value may have any non-negative value, but the number of counts arising from this mean will obviously be integer. According to the Poisson distribution, the probability that this count will be n is

$$p(n) = \frac{e^{-\lambda_k} \lambda_k^{\,n}}{n!} . \qquad (1)$$

Note that n may be zero; no matter how large the value of λ_k, the probability of such an event is $e^{-\lambda k}$. The distribution has only one parameter, λ_k, which is, in fact, both the mean and variance of the distribution. An important property of the distribution is that any one count is statistically independent of all other counts; i.e., the covariance of the j and k^{th} counts is zero when j is not equal to k. This permits easy synthesis of Poisson count data and also facilitates theoretical calculations of test statistic distributions.

Two obvious mechanisms or tests for detecting a "signal" in a time series are to find the largest single count in the dataset (peak detection) or to find the first count in the dataset to exceed a given value (threshold detection). In the former case the location of interest is assigned to the center of the time bin containing the selected peak count; in the latter case, the result could simply be chosen to be the center of the first time bin which exceeds the threshold level, or it could be located more precisely using a linear interpolation based on the fractional value by which the threshold was exceeded above the preceding bin. While significant modifications to these methods will later be examined to improve the location evaluation, these two approaches form the basis of all the criteria to be considered.

It is instructive to derive the probability distributions for the counts selected by these two criteria, since they highlight the difficulties inherent in the estimation problem. First, consider the probability distribution of the peak count. In our treatment, we have chosen to select the peak to be the first of any number of equal largest counts, i.e., the peak count exceeds all preceding counts and equals or exceeds all succeeding counts. Consider now the k^{th} of N counts, and let the expectation of this count be λ_k. The probability that this count attains the value i is

$$p_k(i) = \frac{e^{-\lambda_k} \lambda_k{}^i}{i!}, \qquad (2)$$

and the probability that this count will be less than or equal to i is

$$P_k(i) = \sum_{j=0}^{i} p_k(i) \, . \tag{3}$$

If the k^{th} count is i, the probability that this is the largest count is

$$P_{pk}(i) = P_1(i-1)\, P_2(i-1) \ldots\ldots P_{k-1}(i-1)\, P_{k+1}(i) \ldots P_N(i)$$

$$= \prod_{j=1}^{k-1} P_j(i-1) \prod_{j=k+1}^{N} P_j(i) \, . \tag{4}$$

In evaluating the overall probability that the k^{th} count will be designated as the peak, this term must be weighted by $p_k(i)$, the probability that the k^{th} count will be i. Thus the probability that the k^{th} count will be designated to be the peak count is

$$P_{pk,k} = \sum_{i=1}^{\infty} p_k(i) \prod_{j=1}^{k-1} P_j(i-1) \prod_{j=k+1}^{N} P_j(i) \, . \tag{5}$$

For the first count,

$$P_{pk,1} = \sum_{i=0}^{\infty} p_1(i) \prod_{j=2}^{N} P_j(i) \, . \tag{6}$$

By a similar argument we can show that if K is the threshold to be exceeded, the probability that the k^{th} count will be the designated count is

$$P_{T,k}(K) = (1 - P_k(K)) \prod_{j=1}^{k-1} P_j(K) . \qquad (7)$$

For the first count,

$$P_{T,1}(K) = 1-P_1(K) . \qquad (8)$$

Note that a sum over k of Equation (7) represents the probability that the threshold will occur at all in the dataset; this is the complement of the probability that all counts will be less than or equal to K:

$$P_T(K) = \sum_{k=1}^{N} P_{T,k}(K) = 1 - \prod_{k=1}^{N} P_k(K) . \qquad (9)$$

Formulae (5) through (9) may appear to be formidable, but the calculation is trivial in the computer. We have taken advantage of the fact that the cumulative probability terms $P_k(K)$ are also required in the simulation (which is discussed in the next section) and have coded these formulae as a front-end package to the simulator; this approach provides us with an independent calculation capability which acts as a built-in check on the simulator performance.

DESCRIPTION OF THE SIMULATION

Concept

The simulation process consists first of generating a large number (typically 100 or more) of digitized pulse waveforms, or "data sets," with Poisson distribution about a specified mean signal. A given set of pulse location estimation "procedures" are then applied to each data set. For the laser bathymetry application, eight procedures have been applied; these consist of center of peak bin, centroid of a specified band of bins surrounding the peak, and frontwards- and backwards-looking proportional thresholds at the 20%, 50%, and 80% levels (illustrated in Fig. 1). The mean location, standard deviation, standard error in the mean, and success probability (fractional number of location determinations compared to the number of attempts) for the ensemble of data sets are calculated for each procedure. A bias may be obtained for each procedure by noting the difference between the calculated mean location and the "true" or expected location.

The simulation is performed over the ensemble of data sets to determine pulse location statistics for a single pulse, but a bathymetric measurement is based on two independent pulses: the surface return and the bottom return. The "precision" or standard deviation of a simulated bathymetric measurement is then the root sum of squares (RSS) of the standard deviations for two independent pulses, and biases are summed algebraically to obtain the bathymetric bias for the procedure. The probability of a successful depth measurement is the product of the individual success probabilities.

Data Synthesis

Consider a continuous waveform representing an analog pulse superimposed arbitrarily on a time axis which has been divided into a number of adjacent "increments" or time slices of equal width. In each increment, or "bin", the analog signal is integrated and normalized to yield a single value which is displayed as a constant output level for the duration of that bin. The shape of this digital representation will depend strongly on the location of the input rate peak with respect to the edges of the time bins as seen in Fig. 2.

It is clear that if random noise is added to the pulse, a peak location solidly in the center of a time bin will generally remain in that bin, but a peak location near the edge between two bins will cause the sampled peak to jump back and forth between the two. The standard deviation of the estimate of a pulse location procedure can rightfully be expected to be larger for the latter case -- by an amount which depends on the procedure, the pulse width, and the bin width.

In an operational situation, the actual pulse locations may generally be assumed to be uniformly and randomly distributed over the space from the bin edge to bin center. Simulations are performed in which mean rate input pulses are placed at a number of equally spaced locations across the width of a bin and the output standard deviation of the location estimate is averaged over all cases, as follows. If the mean pulse location (for a given algorithm) for the i^{th} peak location across the bin is M_i and the variance about that mean is V_i, the overall variance for peak locations uniformly distributed across the width of a bin is

$$V = \frac{1}{r} \sum_{i=1}^{r} (V_i + M_i^2) - \frac{1}{r^2} \left(\sum_{i=1}^{r} M_i \right)^2 , \qquad (10)$$

where r is the number of positions sampled. Simulations have been performed for r = 10.

The first step in the simulation is the "digitization" of a selected mean input pulse, f(t), into half-bin sized increments. This produces a set, δ_k, of expected (i.e., mean) counts in the given (half-bin) observation times where

$$\delta_k = \int_{t_{k-1}}^{t_k} f(t)\, dt \qquad (11)$$

is the integral of f(t) over the k^{th} time increment. The values, δ_k, in these half bins are then summed into full bins in two ways--with the half bin to the right for one "system" and with the half bin to the left for a second "system." This produces one system with bin values, $\lambda_{k,1}$, in which the impulse peak is at a given location in a bin and a second system with bin values, $\lambda_{k,2}$, in which the peak is a half bin displaced. By running five such sets of peak locations, results are calculated for ten peak locations across a bin.

A set of cumulative Poisson probability tables, one for each increment of each system, is needed as a basis for the construction of the data sets. For mean rates less than 70, the individual probability densities $p_k(n)$ of obtaining a count n from a Poisson distribution with a mean count of λ_k are calculated from Eq. (2) with the iterative algorithm

$$p_k(n) = \frac{p(n-1) \cdot \lambda_k}{n}, \qquad (12)$$

where $p_k(0) = e^{-\lambda_k}$.

For mean rate λ_k greater than 70, the algorithm

$$\ell n \left\{ p_k(n) \right\} = \ell n \left\{ p(n-1) \right\} + \ell n(\lambda_k/n) \quad (n > 0)$$
$$\ell n \left\{ p_k(0) \right\} = -\lambda_k \tag{13}$$

is invoked to avoid machine underflows. We have found it necessary to compute the terms of the various algorithms using at least 10 significant digits; truncation problems can otherwise occur (especially at the higher mean counts).

Cumulative probability tables are then generated from the probability density functions. For values of λ_k less than 100, the cumulative probability table is constructed from 0 through $\lambda_k + 10\sqrt{\lambda_k}$; for larger values of λ_k the table is constructed for integers between $\lambda_k - 7\sqrt{\lambda_k}$ and $\lambda_k + 7\sqrt{\lambda_k}$ with the assumption that any integers occurring outside these limits can be assigned to the limits themselves without significantly affecting the probabilities. The resulting cumulative distribution function (CDF) tables are stored in a linear array with the starting locations of the tables for each increment stored in a separate array. This procedure provides two important advantages. First, it minimizes the storage requirement by storing only the relevant elements; and second, it provides significantly faster access and operation than the multi-dimensional format.

Once the CDF tables are generated, a count n_{kj} for the k^{th} increment of the j^{th} dataset for each system is computed as follows. A random number, ρ_{kj}, with a uniform distribution between zero and one is generated, and the CDF tables are searched to determine the minimum integer n_{kj} such that

$$P_{kj}(n_{kj}) \geq \rho_{kj} , \tag{14}$$

where P_{kj} (n_{kj}) is the tabulated cumulative probability of obtaining a Poisson integer less than or equal to n_{kj}. The search for n_{kj} is initiated at the most probable region, λ_k, and the search step adjusted to optimize a two-transit system; i.e., a transit in one direction until n_{kj} is passed, and then a return to its value with a unit step. The required simulation time increases only weakly ($\lambda_k^{1/4}$) with increasing mean rate. Typically, we have been able to simulate and test 1000 sets of 40 increments within about one minute on an IBM 360/91 computer.

The optional selection of Gaussian statistics instead of Poisson statistics in dataset generation has been included in the code for its ability to save computer time when large mean rates are involved. Integers with a Gaussian distribution can be generated directly (i.e., without constructing and sampling CDF tables) by invoking the Central Limit Theorem which says that the distribution of the sums of sets of random numbers from any distribution is nearly Gaussian. In this simulation sums of 12 uniformly distributed random numbers on (0,1) are used to produce Gaussian integers of the form N(6,1)--a Normal or Gaussian distribution with a mean of 6 and a standard deviation of 1. These integers are then normalized to the range required for the appropriate λ_k. This option is invoked only for peak mean rates greater than about 30.

While the simulator possesses the capability to generate a broad range of pulse shapes, the specific mean rate distributions, f(t), considered here are generally of Gaussian character (recall Fig. 1). (This temporal pulse shape is completely different from and should not be confused with the probability distribution involved.) The asymmetric shapes utilized to simulate arriving laser pulses consist of a Gaussian leading edge of a given width combined with a Gaussian trailing edge with a different (longer) width. (Triangle shaped pulses have also been examined, and the results are very similar.) Pulse shape and width will henceforth be denoted by listing the one standard deviation (1σ) leading- and trailing-edge widths (in nanoseconds) separated by a colon, i.e., 3:5 or 5:20.

All pulse location procedures to be tested are applied to the same ensemble of data sets; this insures that observed differences are due to the procedures and not statistical anomalies in data set generation. Random errors in the output statistics for pulse location procedures due to the particular random number sequences involved in data set generation can be constrained to any desired level by selection of a sufficiently large number of data sets per ensemble. A minimum value of 100 data sets per ensemble was used; this provides an expected standard error in the estimate of the standard deviation of pulse location (for a given procedure) of $1/\sqrt{2n} = 1/\sqrt{200} \sim 7\%$.

Bathymetric lidar is intended for daytime use as well as night. Under daylight conditions, the lidar return pulses are superimposed on the additive background level of the reflected solar background. In the simulator, a selectable constant mean background rate is added to all bins in the mean rate input pulse data prior to generation of the probability tables and data sets. This mean rate, having exerted its influence on the distribution of Poisson counts, is then subtracted from all data sets prior to pulse detection/location in order to simulate a realistic system in which this would be accomplished (either by ac-coupling or digital data processing) to remove any pulse location bias which could be caused by an uncorrected background level. Negative values generated by this subtraction are set to zero.

The gain of the system and the number of bits in the digitizer act to truncate certain ranges in the number of photoevents into a discrete "count." For example, if 10 photoevents are required to produce a single count, then both 21 and 29 photoevents will produce 2 counts, while 30 photoevents will produce 3 counts. This truncation effect can alter counting statistics, and hence the number of photoevents per count is included in the simulation as an input variable.

Pulse Location Estimators

The first step in pulse location is pulse detection. In the simulation, a pulse is "detected" at the site of the maximum integer in the data set or the first of equal maxima. The location and magnitude of the detected peak are stored for use in the pulse location procedures. Three basic types of pulse location procedures have been analyzed: these are center of peak bin, centroid of a region linked to the peak bin, and threshold.

If we were to base our pulse location estimation on only a selected count, i.e., the peak, we would have to assign the pulse location to an arbitrary point within the corresponding time bin. With no further information, we could do no better than choose the middle of the time bin to be the required location. In reality, pulse peaks will be distributed uniformly across the bin width. The minimum RMS error for a peak detector is thus the bin width divided by the square root of twelve (the standard deviation of a uniform distribution). A procedure which allows a location estimate anywhere within a bin would therefore be preferable. With little a priori knowledge of the return pulse shape, curve fitting estimators appear to be overambitious; we have thus considered centroid and threshold estimators.

Centroid based locators are defined by choices of the length of the centroid region, whether the length is fixed or variable (depending on the pulse width), and of the relative location of the peak within the region (i.e., a centroid of only the leading edge, one balanced symmetrically or asymmetrically on the peak, or one encompassing the entire pulse energy). Results have been determined for slightly asymmetric centroids (matched to pulse shape) of various lengths "centered" on the peak bin.

Threshold based locators can be categorized as fixed level or fractional (to peak height), and they can be frontward searching or backward searching (in time). The performance of any of these variations depends strongly on the specific threshold level selected. In the simulation, the exact threshold location is selected by identifying the bin in which the specified level is passed and interpolating to a fractional location between that and the previous bin based on the amplitudes involved.

RESULTS

Prediction of AOL Performance

This pulse location estimation (PLE) simulator was originally conceived for the purpose of predicting the limiting precision (random depth measurement error as a function of bottom return signal strength) imposed by the hardware in the NASA/AVCO Airborne Oceanographic Lidar (AOL) system which was being test flown for NOAA in the bathymetry mode. Without this information, there would have been no way of allocating the measured random errors between basic design constraints and unknown causes which would require further investigation. As seen in Fig. 3, the system precision (for calm sea conditions) and the simulation results are in excellent agreement. Not only did this give us confidence that the AOL was performing to its design limits, but it also, in turn, verified the performance of the simulation for further predictive purposes.

The simulation also produced a rather surprising result: fractional threshold algorithms (among others) applied to single asymmetric pulses yield offsets (compared to the "true" location) toward the tail of the pulse with magnitudes as large as 30 cm for the cases studied. It is important to note, however, that, in a two pulse measurement, these offsets will be self cancelling to the extent that the two pulses have the same shape and duration. For a "stretched" bottom return, a net depth measurement bias in the deep direction will result as the difference between the offsets for the two single pulse cases. The offsets (and hence resultant bias) are very small for small (20%) threshold fractions and can become large for high (80%) threshold fractions. For example, given a 2.5 ns bin width and a 20 photoelectrons per nanosecond (pe/ns) peak rate, a pulse stretching from 3:5 ns to 5:20 ns would exhibit a 5 cm bias with a 20% threshold, a 6 cm bias with a 50% threshold, and a 13 cm bias with an 80% threshold.

Sensitivity Studies

In a shot-noise limited system, the performance (measurement accuracy) depends not only on the return signal strength to ambient background ratio; but on the absolute magnitudes of these two individual components. Simulations have been conducted for two diverse conditions: low absolute rates (consistent with AOL performance) and high absolute rates (consistent with the design parameters of the Hydrographic Airborne Laser Sounder (HALS) presently being designed by AVCO for the U.S. Navy). The distinction arises not from different environmental conditions (such as night versus day), but rather from different tranceiver configurations (output power, optics, etc.).

Low Signal Rates

When photon arrival rates are low, counting statistics and the resulting shot-noise level are very sensitive to factors which cause changes in the rates. Such factors as pulse width, integration time (bin width), and amplifier gain (truncation in digitization) are important because they will have larger effects here than in high rate systems. Simulations were performed for all combinations of the parameter sets listed in Table 1.

Table 1. Simulated Parameters and Procedures

Pulse width:	1:2 ns, 3:5 ns, 4:10 ns, 5:20 ns (as defined in Fig. 1)
Bin width:	1.5 ns, 2.5 ns, 5.0 ns
Algorithms:	centroid (a 6 bin window with the peak in bin 3 denoted "6C3"), center of peak count (denoted "PK"), and frontward- and backward-looking fractional thresholds at 20%, 50%, and 80% of the peak amplitude (denoted F20, F50, F80, B20, B50, and B80).

For the AOL case, the mean peak rate was varied from 2 photoelectrons per nanosecond (pe/ns) to 20 pe/ns in steps of 2 pe/ns. The solar background for daytime operation was estimated to be 8 pe/ns, while the dark current of the PMT was estimated to be 2 pe/ns for nighttime operation. Truncation levels of 1, 4, and 8 pe/"discrete" count were exercised. Sample results appear in Figs. 4 through 18.

The effect of the pulse location algorithm is examined for the case of a typical unstretched pulse at night in Figs. 4 and 5 which present the bias in mean pulse location and the standard deviation about that mean (as function of the peak signal strength) for various algorithms. (Notation used in identifying the algorithms on the plots is listed in Table 1).

It can be seen in Fig. 4 that as the peak signal strength is reduced, the mean pulse location remains constant and stable for all algorithms except F20 until about 6 pe/ns (a peak signal to mean background ratio of S/B=3). The F20 result (i.e., for a forward-looking 20% threshold searching from the beginning of the data set) becomes unstable for peak signals of less than 12 pe/ns. This is indicated by the drop in the mean F20 pulse location below its (correct) high signal strength value. This drop is caused by shot-noise induced false early detections generated in the region between the start of the data set and the true 20% threshold location.

This results, as clearly evidenced in Fig. 5, in a very large standard deviation for the F20 algorithm below a peak signal strength of 10 pe/ns. The remaining algorithms produce pulse location precisions (with magnitudes related to the algorithm) which do not increase significantly until the peak signal rate drops below about 8 pe/ns (S/B=4). It can be seen that, for this parameter set, the best performance (lowest asymptotic standard deviation at high peak rates) from a threshold detector is about 6 cm at 20 pe/ns derived from B20, F50, and B50; while F80 and B80 are slightly noisier at 9 cm; and peak detection at 13 cm is much noisier and yields a result twice as large as that for the preferred algorithms. The 6-point asymmetric centroid is, in this case, well matched to the pulse and hence offers the lowest standard

deviation of 4 cm. This will not always be the result, however, because the precision of a centroid-based pulse location estimate depends strongly on the "matching" of the size and location of the centroid window to the given pulse. The high standard deviation associated with the peak location is, however, as will be seen, a general result.

A similar set of results is presented in Figs. 6 and 7 for a stretched (e.g., by underwater propagation) return pulse. It is seen in Fig. 6 that the results for the means are similar to but subtly different from those for the previous "nominal" pulse. The peak location is biased above its actual location by the high probability of detecting a peak (caused by noise spikes) on the long, slowly-decaying trailing edge of the pulse. As the peak signal rate is reduced, several of the fractional threshold derived means rise (rather than fall as in the previous example) for the same basic reasons. The F20 algorithm again becomes unstable below 6 pe/ns due to early detections in the noise preceding the pulse.

Most precision results for this case, as seen in Fig. 7, are qualitatively similar to the former case, but quite different quantitatively. One major qualitative difference is the relative performance of the centroid detector whose standard deviation of 29 cm (at 20 pe/ns) falls far above that for the 20% and 50% thresholds and just below that of the peak detector at 33 cm. This occurs because the 6 bin window is no longer large enough to encompass the entire pulse, and the centroid result "jitters" with the movement of the peak detector to which it is tied. In addition, the B80 at 30 cm is not as good as the F80 at 21 cm due to the flatter shape of the top of the pulse. The F50 and B20 at 10 cm are again preferred; the F20 is lost in the noise at low signal levels, and the B50 is still feeling the effects of the elongated pulse shape (as evidenced by unusually prolonged elevated values at middle-sized peak signal rates).

It is very important to note the effect of the increased pulse width on the limiting precision of the various algorithms. In the field, the bottom return pulse width will increase more than linearly with increasing depth (Guenther and Thomas 1981). It would not, therefore, be appropriate to select an algorithm whose basic precision limitation is strongly sensitive to pulse width. The aforementioned limiting precision results are compiled in Table 2.

Table 2. Effect of Pulse Width on Limiting Precision (at 20 pe/ns) for Various Algorithms with 2.5 ns Bins

Algorithm* †	Limiting Precision (cm)		Increase (cm)
	Pulse Width		
	3:5 ns	5:20 ns	
6C3 centroid	4	29	25
PK (peak)	13.5	33	19
B20 threshold	6	10	4
F50 threshold	6	10	4
F80 threshold	9	21	12
B80 threshold	9	30	21

It is clear that the B20 and F50 algorithms not only produce some of the lowest standard deviations for unstretched pulses but also are the least sensitive to pulse stretching. This is further illustrated in Fig. 8 which compares the performance of a peak detector against F50 for increasing pulse widths.

* for notation see Table 1.

† F20 and B50 are excluded due to poor performance as described in the text.

The choice of the "optimum" threshold fraction (defined in terms of the lowest standard deviation) depends to a certain extent on both pulse width and bin width as seen in Fig. 9. Here we see the value of the mean peak signal strength required to reduce the standard deviation of the pulse location estimate to 10 cm or less as a function of the threshold fraction.

For narrow (1:2 ns) pulses fairly large threshold fractions are preferred. As the pulse width increases to 3:5 ns, a minimum forms at a threshold fraction of about 0.5. As the width increases to 5:20 ns, the required signal becomes larger, and higher threshold fractions become increasingly undesirable--particularly for narrow bin widths. The increase in required signal is due to the decreasing slope at the detection point. The increasingly poor performance for narrow bins is a result of the classic struggle between resolution and accuracy. As the bin size is decreased, the resolution improves, but the encompassed signal decreases, and the counting statistics cause the overall accuracy to degrade. Over a range of pulse widths, it can be seen in Fig. 9 that the optimum threshold fraction lies in the range from 0.3 to 0.6. For the cases illustrated, the wider 5.0 ns bin results in slightly better performance for pulse widths greater than 3:5 ns. This bin size effect will be illustrated again after a few other sensitivities are examined.

Figures 10 and 11 illustrate the effect of raising the background rate to 8 pe/ns to simulate the effect of daytime operation. A comparison of these curves with Figs. 5 and 7 indicates that results for most algorithms are remarkably similar in shape and limiting value (at high signals) and that the curves are basically shifted toward higher peak rates with the "knees" occurring at about 10 pe/ns (P/B $\sim$ 1.2). An exception is F20 which can be seen (in Fig. 10 at the upper right corner and mid-group in Fig. 11) to have been made even worse than before (as might be expected). The effect of the increased (solar) background on the F50 algorithm is seen directly in Fig. 12 for nominal and stretched pulses. For peak signal strengths beyond about 12 pe/ns the differences are reasonably small compared to the desired error budget.

Because of its demonstrated performance, the F50 algorithm will be used as a standard for the remaining sensitivity analyses.

The effect of truncation of counts by the digitizer is demonstrated in Fig. 13 for the F50 algorithm. It is clear that this effect is small compared to many others, and the "standard" value of 4 pe/ns which has been used in previous comparisons will be continued.

The effect of bin size for an F50 algorithm applied to nominal and stretched pulses is seen in Figs. 14 and 15 to be quite small and not worth discussing except to note that for a wide pulse, decreasing the bin width causes a slight increase (rather than decrease) in the random pulse location error--due to the previously mentioned effect of counting statistics.

The effect of pulse width for three bin widths is presented in Figs. 16 through 18. It can be noted that for an F50 algorithm, the reduction in precision due to moderate pulse stretching is only about 3 cm and is not significant.

High Signal Rates

Sample results are illustrated in Figs. 19 through 22. The effect of pulse location algorithm on the behavior of the means and standard deviations of the estimated pulse location is qualitatively similar to that for low signal rates. Specifically, the measurement bias and precision from peak detectors depend strongly on both bin width and pulse width, and the precision is poor (compared to fractional threshold detectors) for either wide bins or wide pulses). For narrow pulses, i.e., 3:5 ns or less, the precision depends on the bin width due to the discrete nature of the detection locations. For a wide (5:20 ns) pulse, the random error dominates due to the high probability of detection along the broad trailing edge, and the standard deviation is thus large and fairly independent of bin widths up to 5.0 ns. A considerable deep bias is also associated with this effect as seen in Fig. 21. Centroid detectors with a fixed number of bins locked to the peak vary considerably in performance depending on relative bin and pulse width and provide tremendous biases for asymmetric stretched pulses (Fig. 21). Both of these detectors are consequently unacceptable.

Fractional threshold pulse locators utilizing linear interpolation between bin centers based on amplitude tend to be fairly well behaved in that most biases and standard deviations are relatively small and exhibit only moderate dependence on bin width and pulse width. As seen in Figs. 19 and 21, 5.0 ns bins can produce fairly large shallow biases for 20% threshold fractions. Biases for 50% and 80% thresholds are typically less than ± 8 cm. The narrow bins are least prone to bias and for most algorithms are biased slightly deep; big bins tend to result in slightly shallow biases. Although the asymptotic standard deviations for all thresholds are reasonably small (Figs. 20 and 22), F50, B50, and B20 provide the best results. F80 and B80 tend to be noisier due to detection at a lower slope (particularly on wide pulses), and F20 is the first to lose precision under low signal-to-noise ratio conditions. B20 performs to low signal rates if a small threshold is desired. For F50, B50, and B20, the bin width effect on precision is quite small.

The most important feature of the precision plots is the signal strength at which the standard deviation rises above some preselected maximum permissible value such as, say, 10 cm. The knees in these high signal rate plots tend to be much sharper than those seen earlier for low rates. It can be seen in Figs. 20 and 22 that for F50 the 10 cm mark is passed at a rate of about 150 or S/B = 0.15. This is a considerably lower ratio than for low rates due to the relatively smaller effect of random noise.

CONCLUSIONS

The effects of pulse shape and duration, integration time, pulse location algorithm, signal level, background level, and digitizer truncation on the pulse location estimation accuracy for quantum limited returns have been studied via a Poisson count simulator utilizing Monte Carlo techniques. Experimental pulse-to-pulse precision data obtained with the NASA/AVCO Airborne Oceanographic Lidar (AOL) system were consistent with simulator predictions for that specific configuration.

The pulse location algorithm is very important in determining both the limiting precision at high signal rates and the signal rate at which the precision drops to an acceptable level. (The latter is also strongly influenced by the absolute peak signal and background rates and must be examined individually for each specific case of interest.) Centroid-type pulse locators require interactive decisions on window size and placement and nevertheless result in large, unacceptable biases on stretched pulses. Correlation or matched-filter locators are not appropriate for application to propagation-stretched pulses. Peak detectors are inherently noisy due both to their discrete nature and to detection in a zero slope region. They exhibit disastrous loss of precision with increasing pulse width and are also prone to a serious deep bias for asymmetric pulses. Fixed level thresholds produce large signal strength dependent biases and are unacceptable. Fractional threshold detectors with linear interpolation between adjacent time bin amplitudes offer the best overall performance. The optimum threshold fraction (of peak height) depends to a certain extent on the other parameters, but of all algorithms examined, the one preferred was the frontward-looking 50% threshold (F50) which offers low bias, low limiting standard deviation, and a rapid approach to limiting standard deviation with increasing signal rates. The backward-looking 20% threshold (B20) is acceptable as long as bin width is maintained below about 3 ns. (Results from the propagation simulation (Guenther and Thomas 1981) indicate that interactive selection of the threshold fraction as a function of nadir angle may provide a tool for constraining the magnitude of propagation-induced

depth measurement biases. Results from this simulation indicate that this can be done without serious loss of precision if bins are kept below 3 ns and low threshold fractions are backward looking.)

For well-chosen fractional threshold pulse locators, such as F50, the effects on precision of digitizer truncation and integration time (bin width) are minimal. The effects of pulse width and shape are small for F50 and, for the simulated parameters, fall within the desired error budget. Specifically, a system with an 8 ns wide laser pulse and integration time between 2.5 and 5.0 ns has the theoretical potential to provide biases and precisions below 10 cm - even though the pulses may be stretched by propagation to 25 ns wide - for reasonable signal to background ratios (which depend strongly on the absolute level of the background). This is advantageous because it implies that there is no need to struggle with state-of-the-art hardware to eke out the last possible nanosecond of performance from either the laser or receiver.

ACKNOWLEDGEMENTS

We wish to thank Lowell R. Goodman whose drive and dedication to airborne laser hydrography helped spur the metamorphosis from technological toy to inescapable destiny. We also thank Rodney McJimpsey who patiently performed much of the tedious data reduction necessary for the presentation of these results.

REFERENCES

Abbot, R.H., and Penny, M.F., 1975: Air Trials of an Experimental Laser Bathymeter, WRE-TN-1509, Weapons Research Establishment, Department of Defense (Australia), Salisbury, South Australia, 28 pp.

Abbot, R.H., Watts, G.J., and Penny, M.F., 1978: WRELADS I Trials Report, ERL-0026-TR, Defense Research Centre Salisbury (Australia).

Cunningham, L.L., 1972: Test Report on Pulsed Light Airborne Depth Sounder (PLADS), Naval Oceanographic Office Technical Note 6620-102-72, U.S. Navy, 53 pp.

Enabnit, D.B., Goodman, L.R., Young, G.K., and Shaughnessy, W.J., 1978: The Cost Effectiveness of Airborne Laser Hydrography. NOAA Technical Memorandum NOS26, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, Rockville, MD., 56 pp.

Enabnit, D.B., and Nield, V.K., 1980: Airborne Laser Hydrography. International Hydrographic Review, Monaco, LVII (2), 93-99.

Goodman, L.R., (ed.), 1975: Laser Hydrography User Requirements Workshop Minutes, January 22-23, 1975, Naational Oceanic and Atmospheric Administration, Rockville, MD., National Aeronautics and Space Administration, Wallops Flight Center, Wallops Island, Virginia, 143 pp.

Goodman, L.R., (ed.) 1976: Laser Hydrography Technical Review Workshop Minutes, August 25-26, 1976, NOAA/National Ocean Survey, Rockville, Md., National Oceanic and Atmospheric Administration, Rockville, Md., 127 pp.

Goodman, L.R., 1978: Laser Hydrography. Proceeding Coastal Mapping Symposium, August 14-16, 1978, Rockville, Md., American Society of Photogrammetry, 105-110.

Guenther, G.C., 1978: Bathymetry Intercomparison: Laser vs. Acoustic. Proceedings Coastal Mapping Symposium, August 14-16, 1978, Rockville, Md., American Society of Photogrammetry, 111-121.

Guenther, G.C., and Goodman, L.R., 1978: Laser Applications for Near-Shore Nautical Charting. Proceedings of SPIE Ocean Optics V, Vol. 160, August 30-31, 1978, San Diego, CA., 174-183.

Guenther, G.C., and Goodman, L.R., (ed.), 1979: Airborne Laser Hydrography Symposium III, October 5-6, 1977, NOAA/National Ocean Survey, Rockville, Md., National Oceanic and Atmospheric Administration, Rockville, Md., 248 pp.

Guenther, G.C., and Thomas, R.W.L., 1980: The Effect of Multiple Scattering in Water on Airborne Lidar Bathymetry Accuracy. Proceedings 10th International Laser Radar Conference, October 6-9, 1980, Silver Spring, Md., American Meteorological Society, 55-56.

Guenther, G.C., and Thomas, R.W.L., 1981: Depth Measurement Biases for an Airborne Laser Bathymeter. Proc. Laser Hydrography Symposium, 30 Sept-3 Oct 1980, Salisbury, South Australia, Defence Research Centre Salisbury, South Australia, in press.

Hickman, G.D., and Hogg, J.E., 1969: Application of an Airborne Pulsed Laser for Near-Shore Bathymetric Measurements. Remote Sens. of Env., 1, 47-58.

Hickman, G.D., Gault, C.S., and Ghovanlou, A.H., 1974: Airborne Laser Shallow Water Bathymetric System. Final Report, ONR/NOAA/USGS Contract No. N000-14-71-C-0202, Sparcom, Inc., Alexandria, Virginia, 41 pp.

Hoge, F.E., Swift, R.N., and Frederick, E.G., 1980: Water Depth Measurement using an Airborne Pulsed Neon Laser System. Appl. Opt., 19, 6, 871-883.

Kim, H.H.,and Ryan, P.T., (ed), 1975: The Use of Lasers for Hydrographic Studies, September 12, 1973, NASA Wallops Flight Center, Wallops Island, Virginia, U.S. National Aeronautics and Space Administration, Washington, D.C., 207 pp.

Kim, H.H., Cervenka, P., and Lankford, C., 1975: Development of an Airborne Laser Bathymeter. NASA Technical Note TND-8079, National Aeronautics and Space Administration, Washington, D.C., 39 pp.

Ryan, J.S., and O'Neil, R.A., 1980: Proceedings of the 19th Annual Canadian Hydrographic Conference, March 20, 1980, Halifax, Nova Scotia, J. Canadian Hydrographers Assn.

Thomas, R.W.L.,and Guenther, G.C., 1979: Theoretical Characterization of Bottom Returns for Bathymetric Lidar. Proceedings of the International Conference on Lasers '78, December 11-15, 1978, Orlando, Fla., Society for Optical and Quantum Electronics, McLean, Va., 48-59.

Van Norden, M.F., 1980: The Hydrographic Airborne Laser Sounder: A Planning and Operational Scenario. Proc. 19th Canadian Hydrographic Conference, March 18-20, 1980, Halifax, Nova Scotia, J. Canadian Hydrographers Assn., 71-79.

Witt, A.K., Shannon, J.G., Rankin, M.B., and Fuchs, L.A., 1976: Air/Underwater Laser Radar Test Results, Analysis, and Performance Predictions. Report No. NADC-76005-20, Naval Air Development Center, Warminster, Pa., 293 pp. (CONFIDENTIAL)

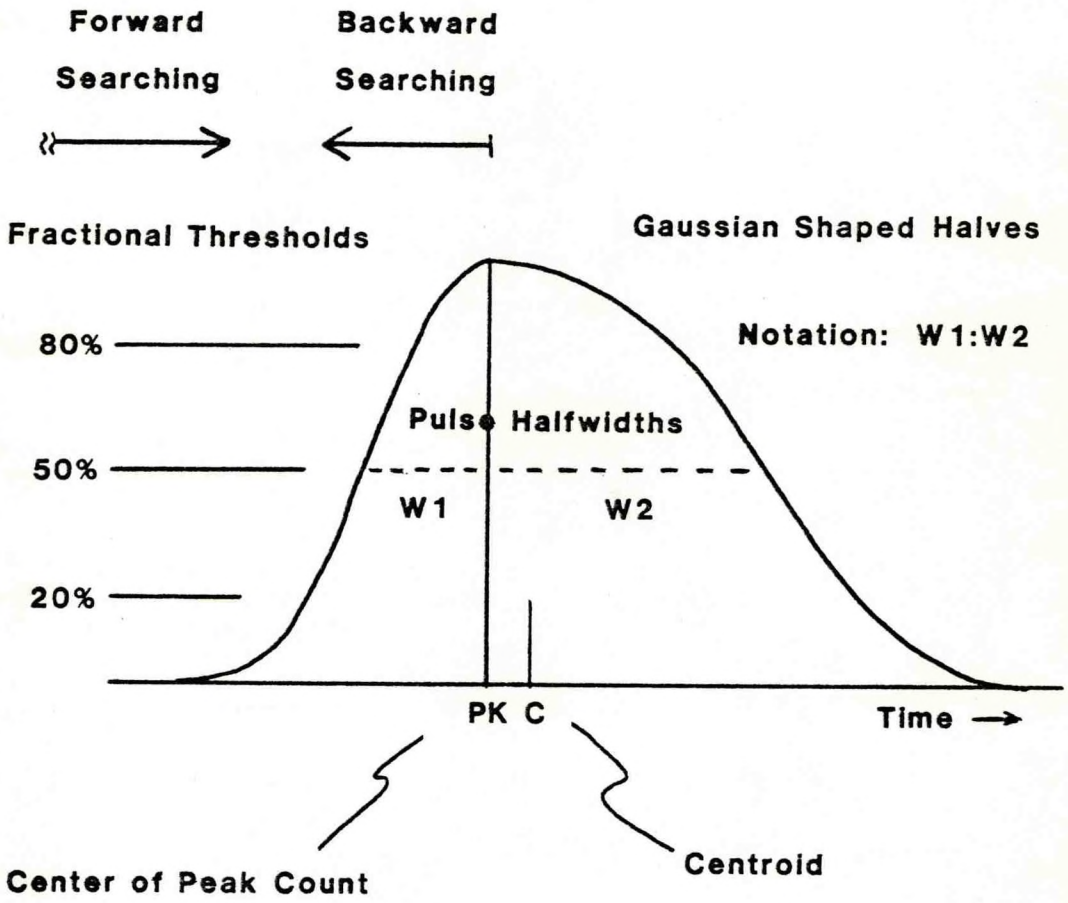

Figure 1. Input Pulse Shape and Pulse Location Algorithms

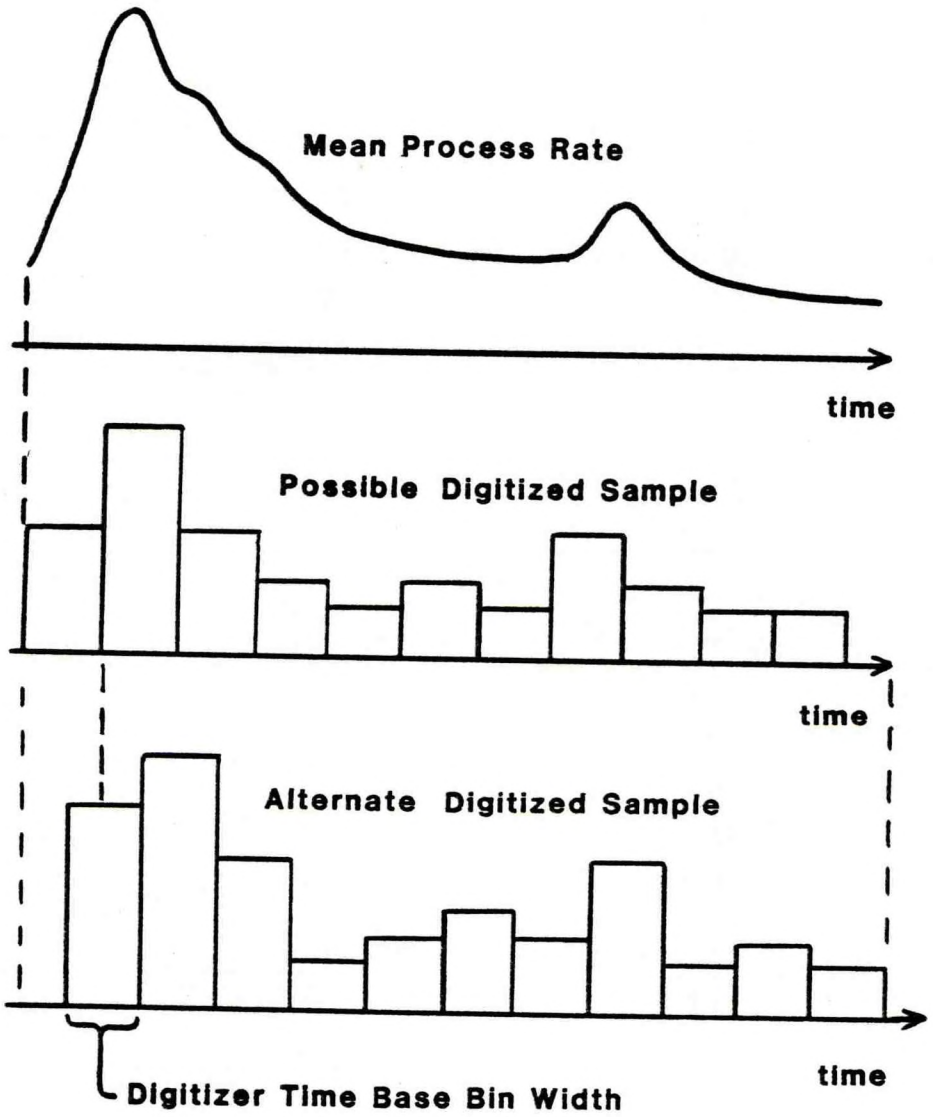

Figure 2. Illustration of Digitization Timing

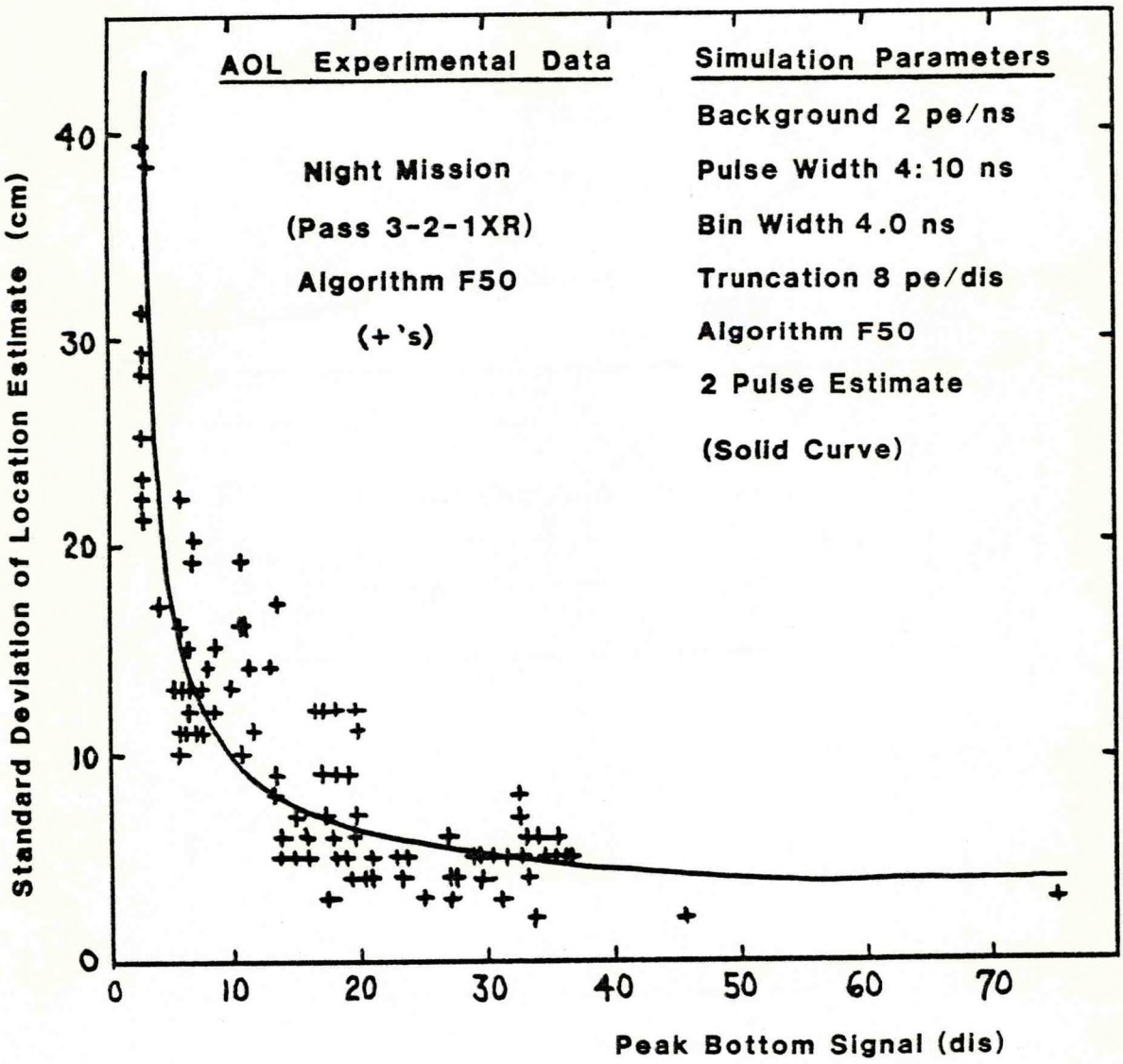

Figure 3. Comparison of Experimental Data with Simulation Prediction

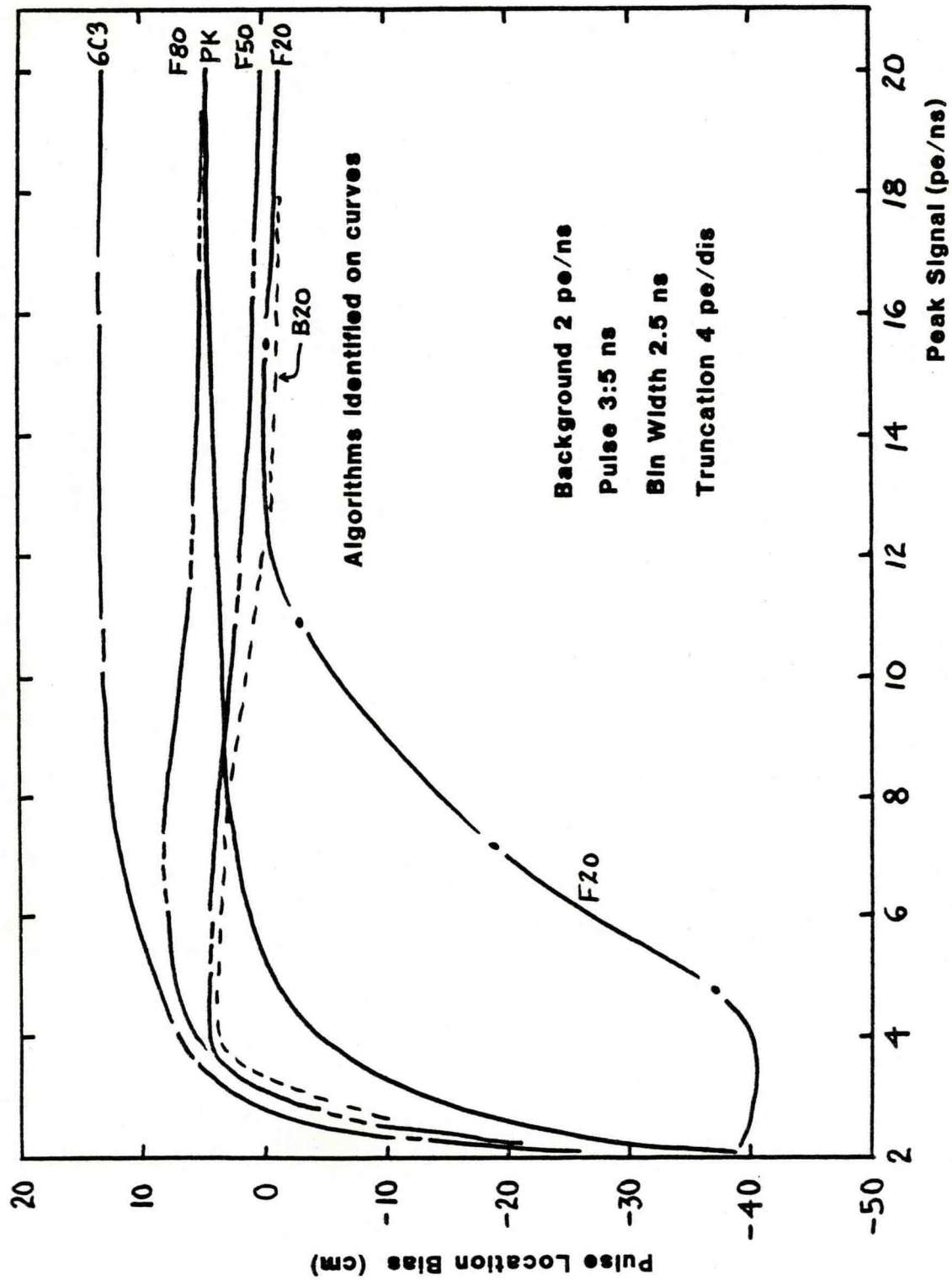

Figure 4. Biases for Nominal Pulse at Night

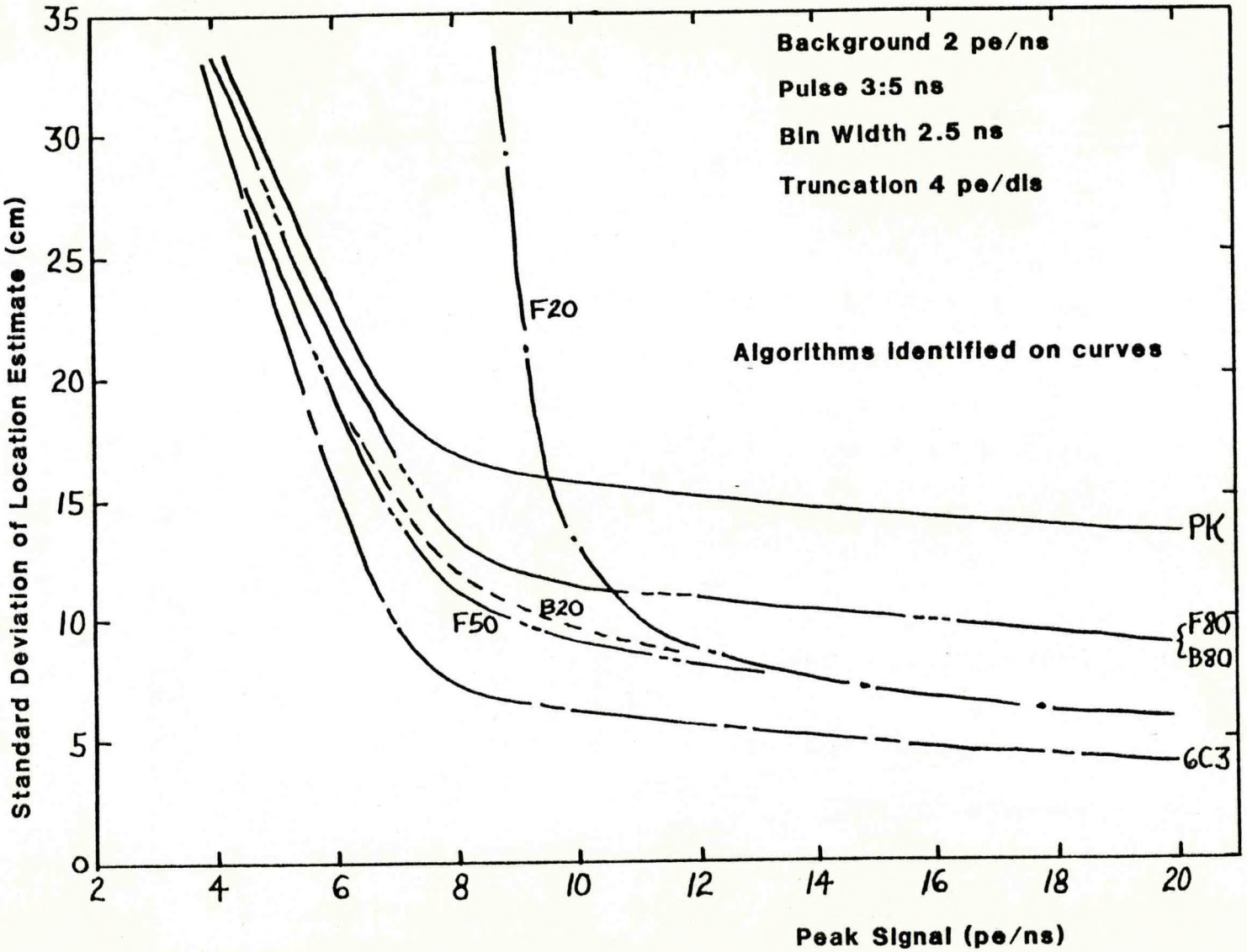

Figure 5. Precision for Nominal Pulse at Night

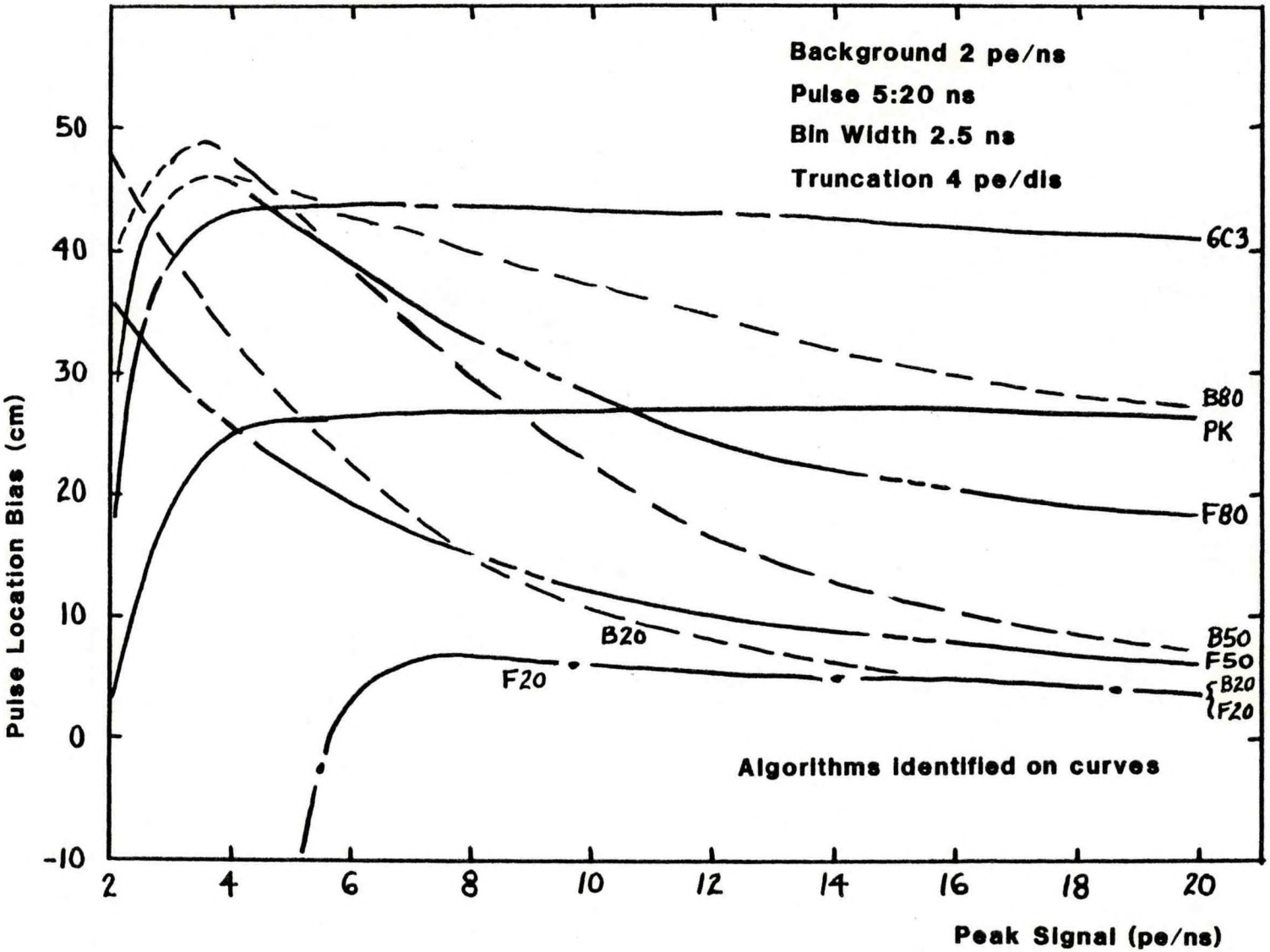

Figure 6. Biases for Stretched Pulse at Night

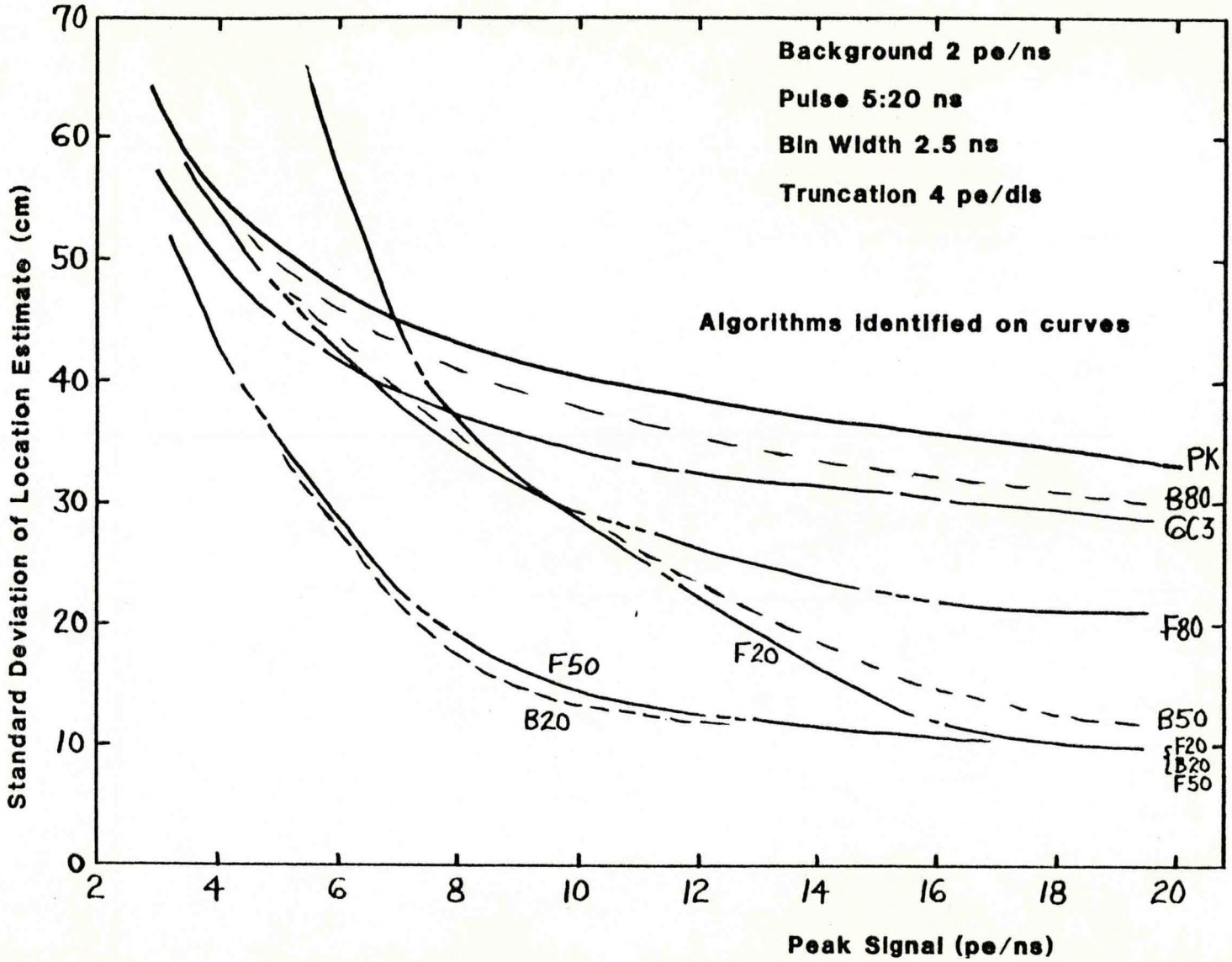

Figure 7. Precision for Stretched Pulse at Night

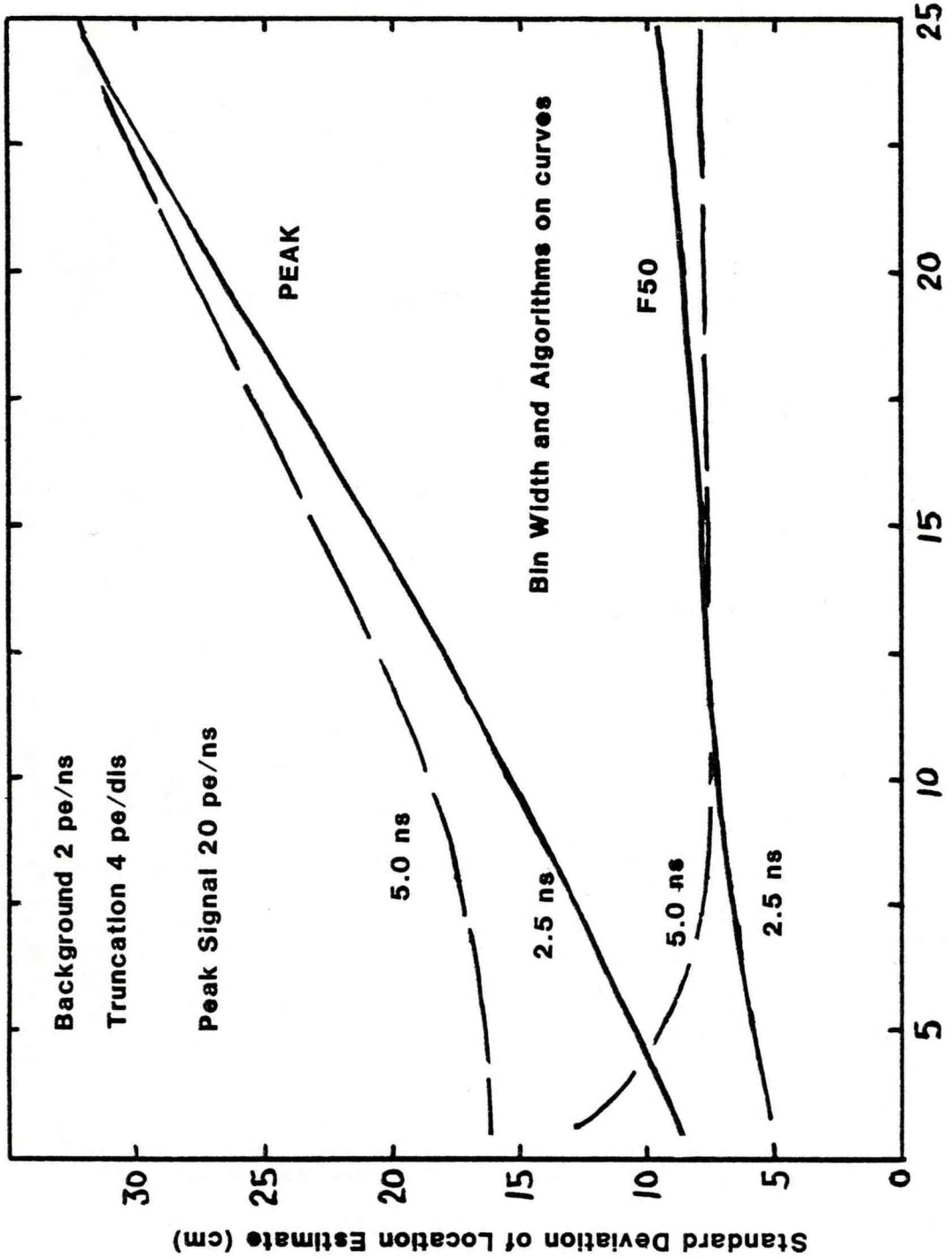

Figure 8. Effect of Pulse Stretching for Two Algorithms

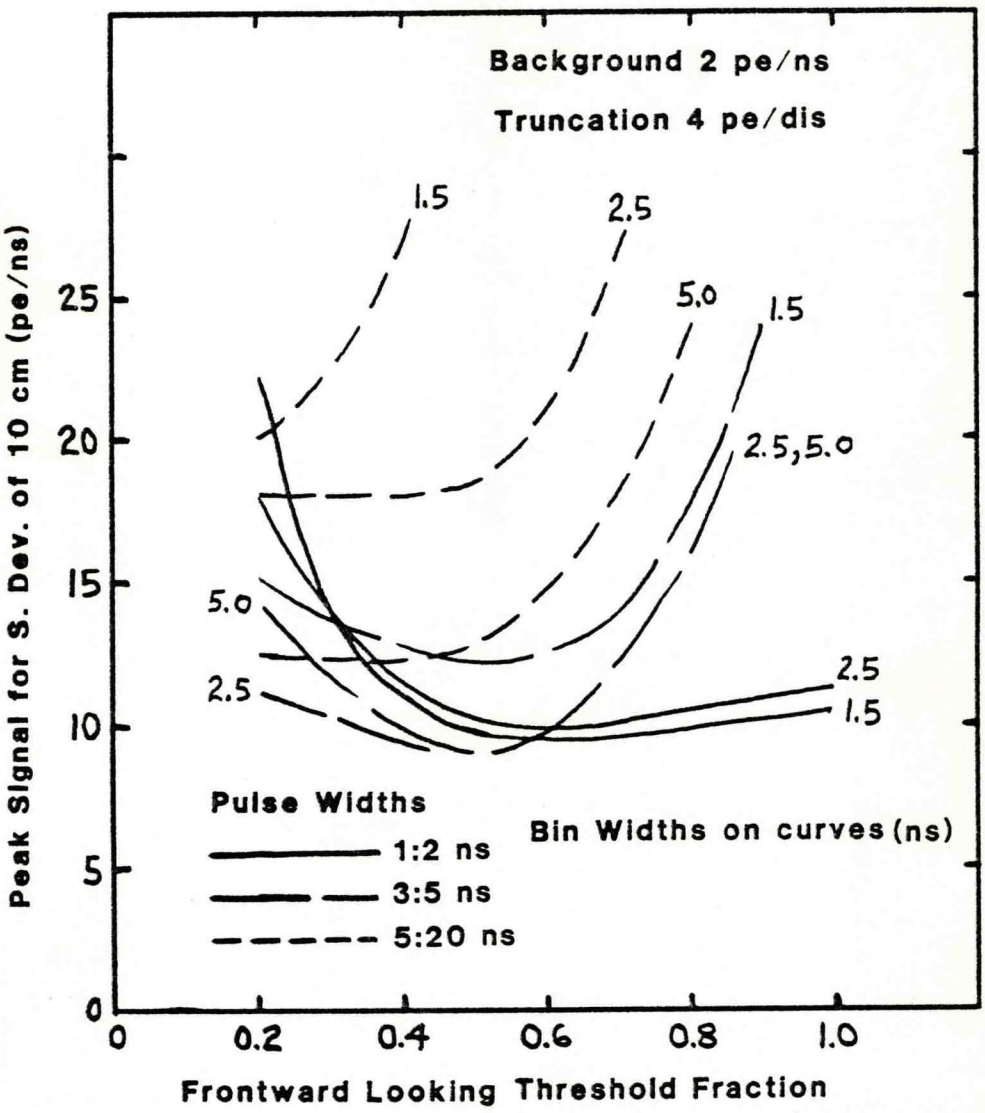

Figure 9. Effect of Threshold Fraction on Signal Required for 10 cm Precision

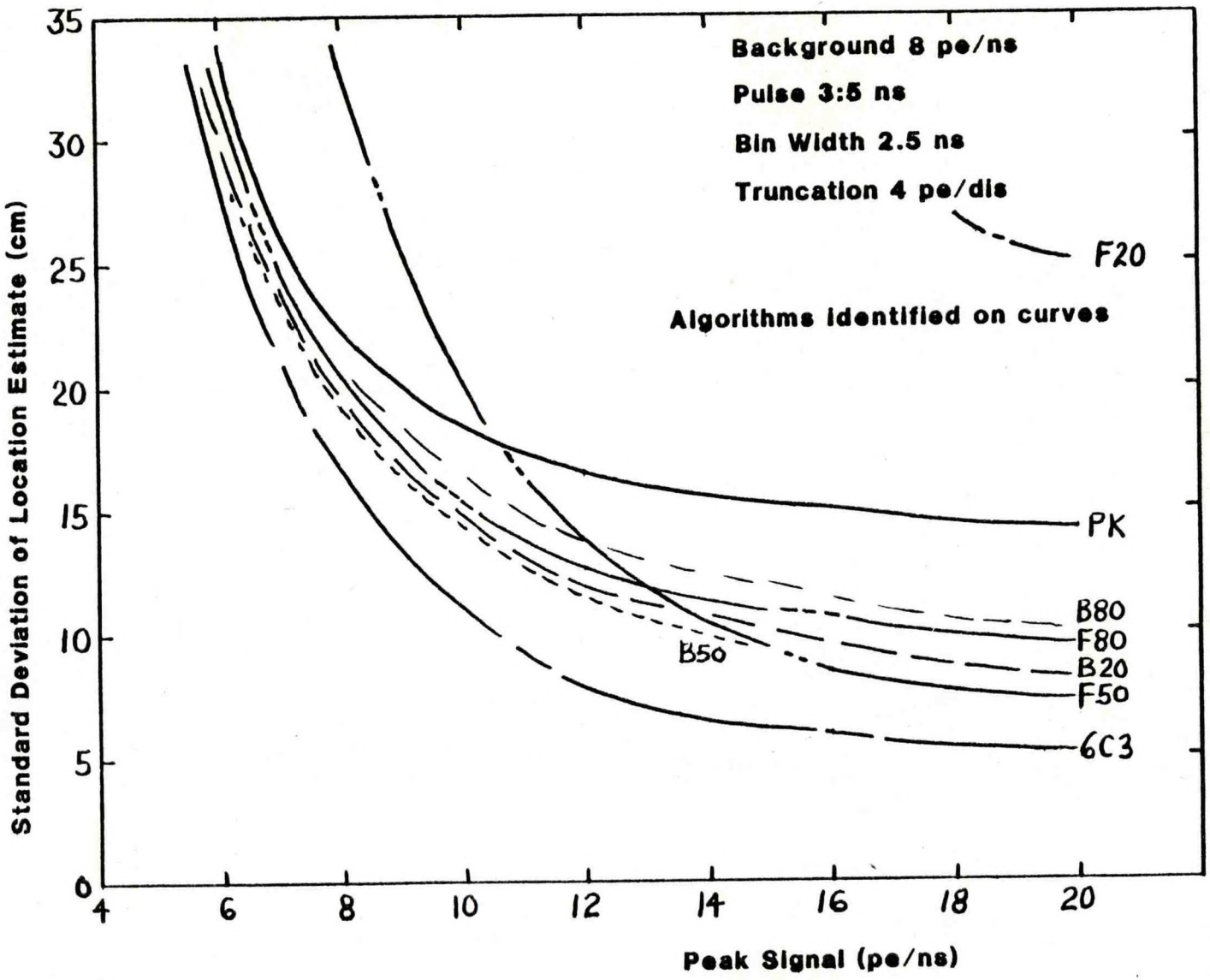

Figure 10. Precision for Nominal Pulse in Daylight

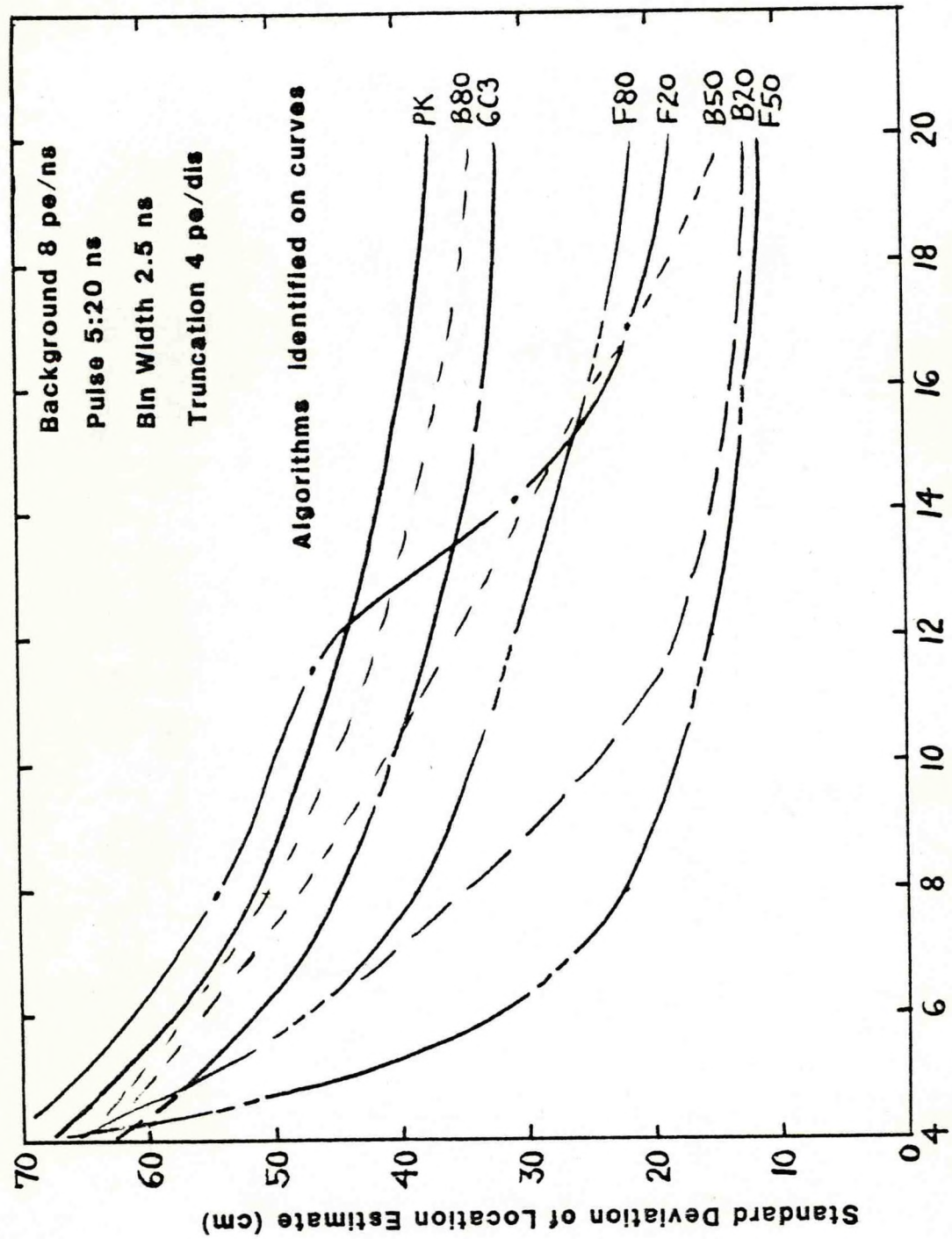

Figure 11. Precision for Stretched Pulse in Daylight

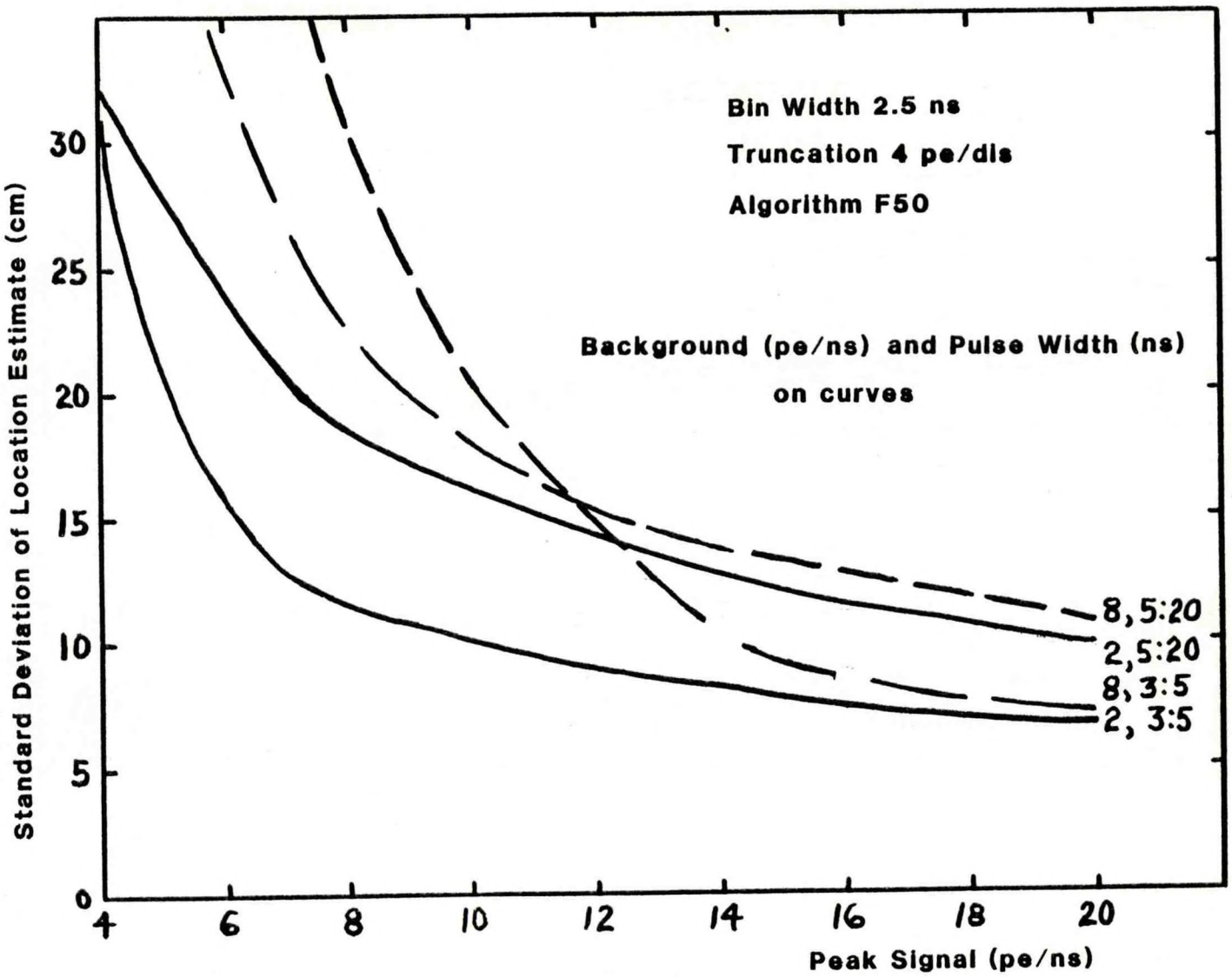

Figure 12. Day vs. Night Effect

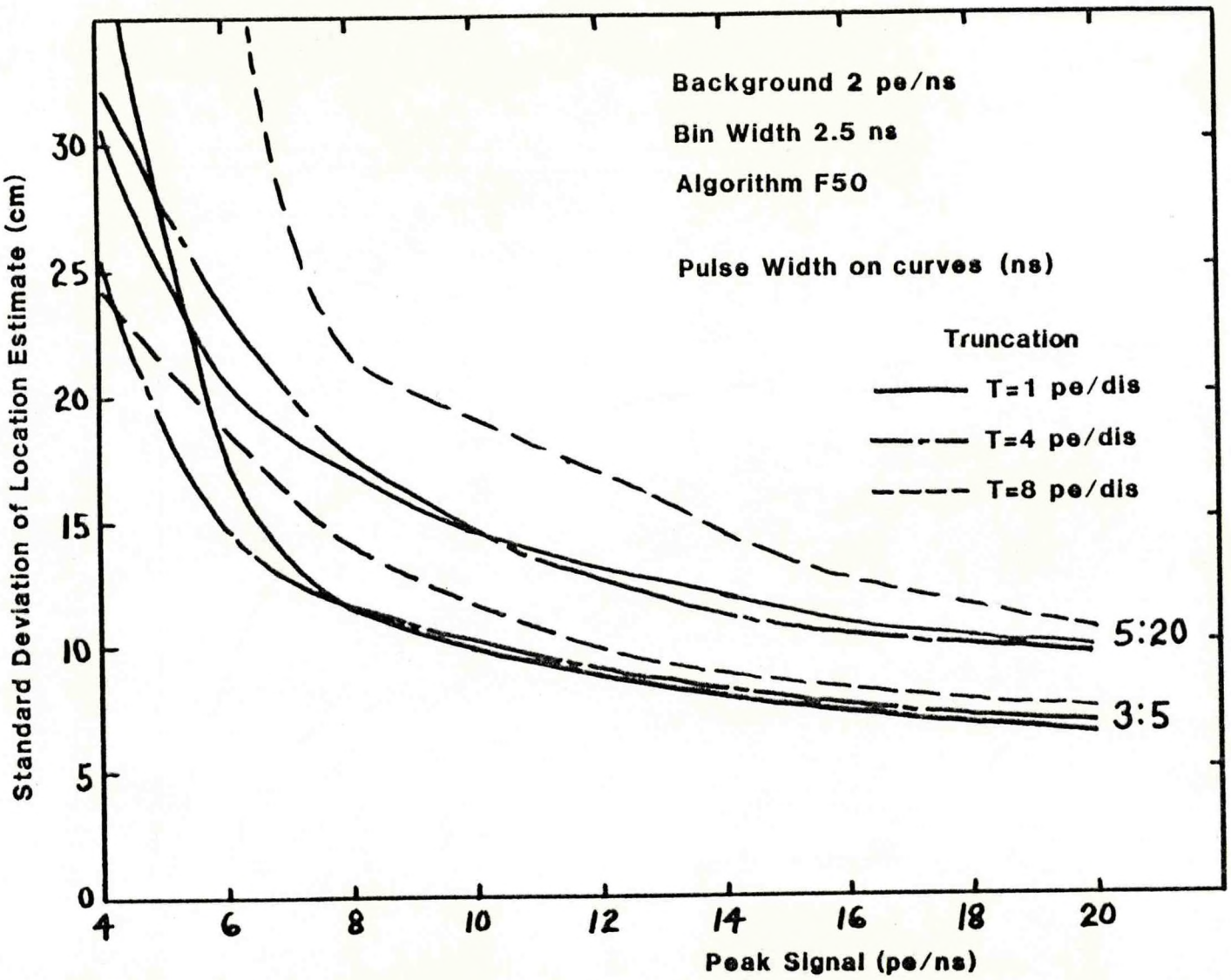

Figure 13. Truncation Effect

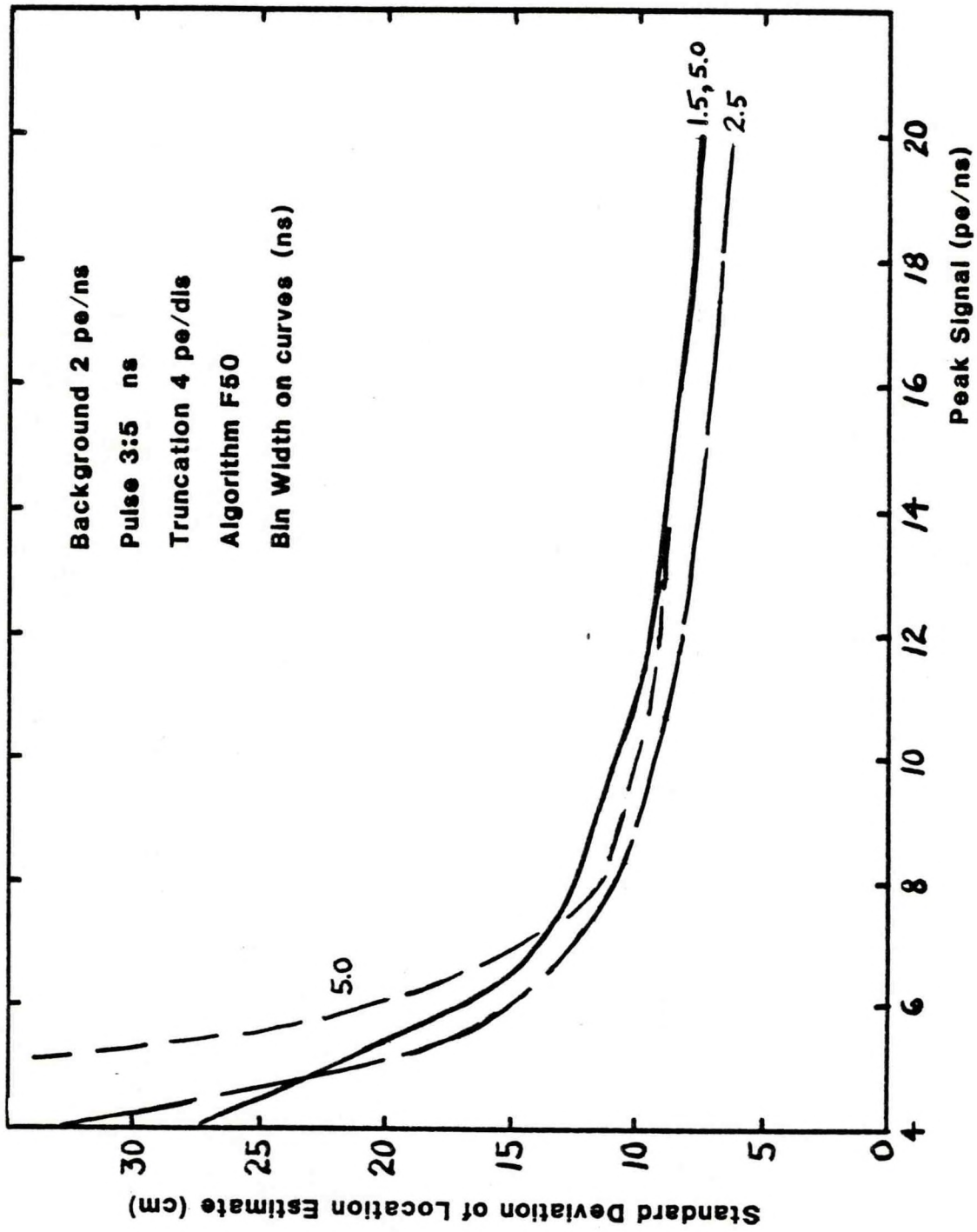

Figure 14. Bin Width Effect for Nominal Pulse

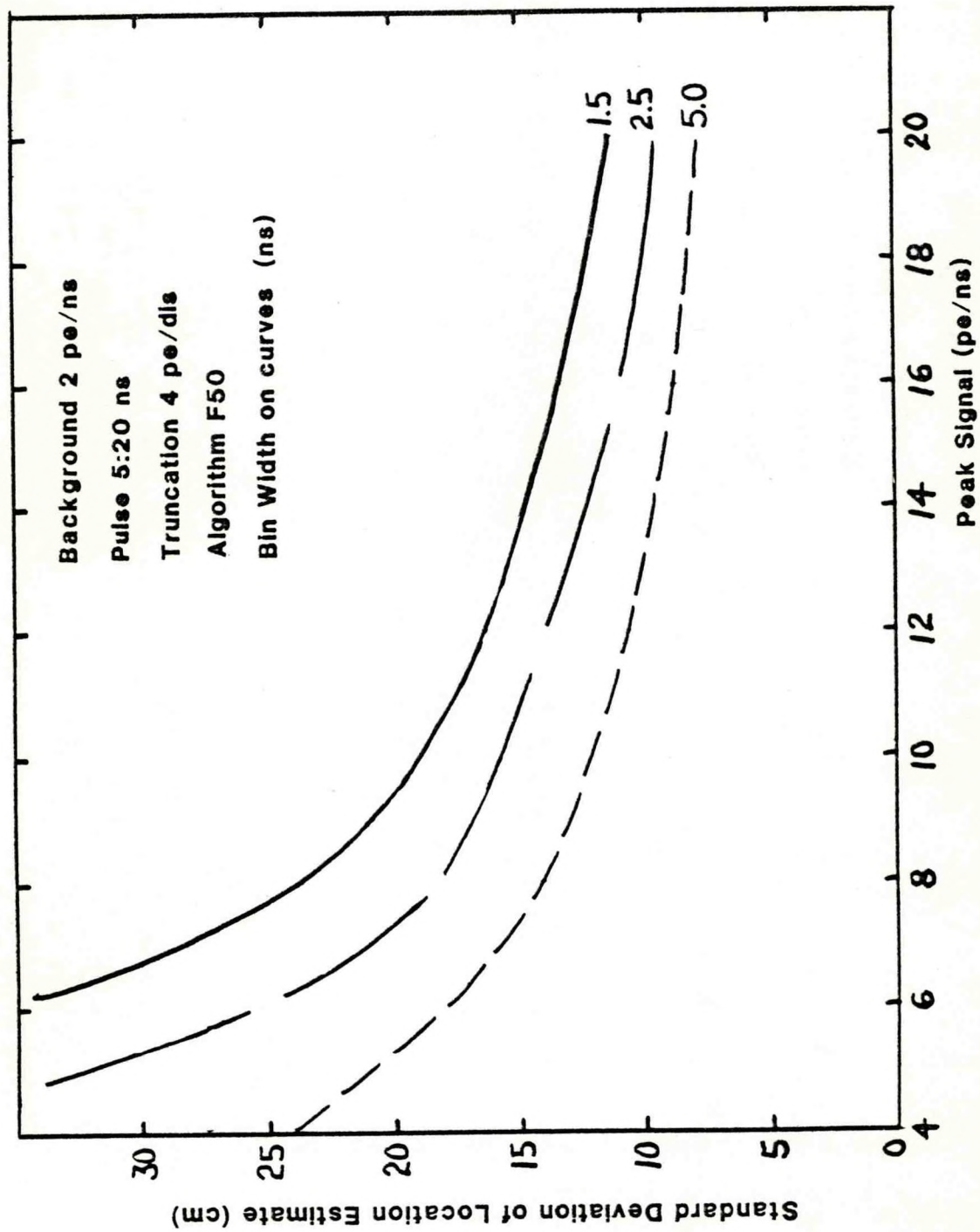

Figure 15. Bin Width Effect for Stretched Pulse

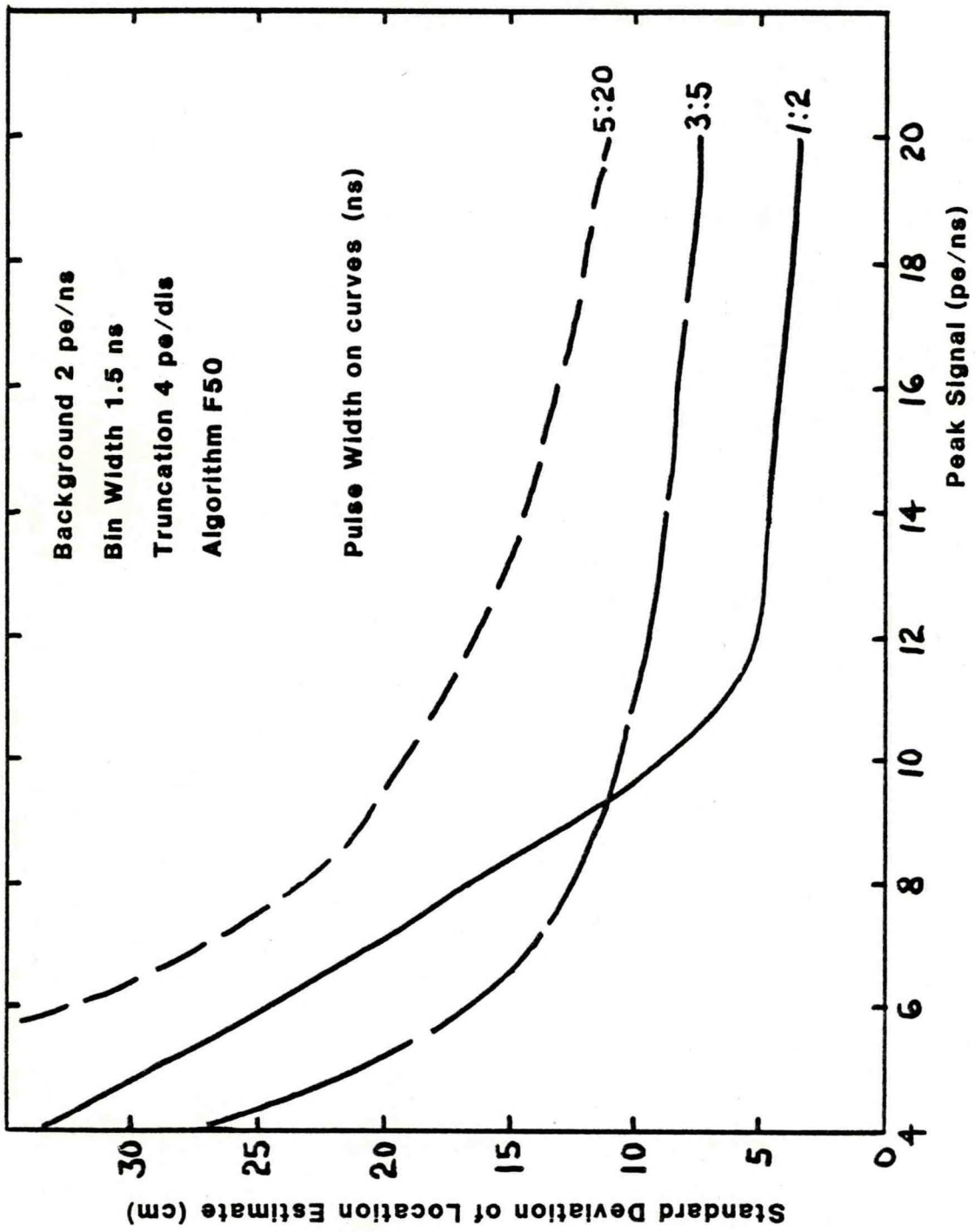

Figure 16. Pulse Width Effect for Small Bin

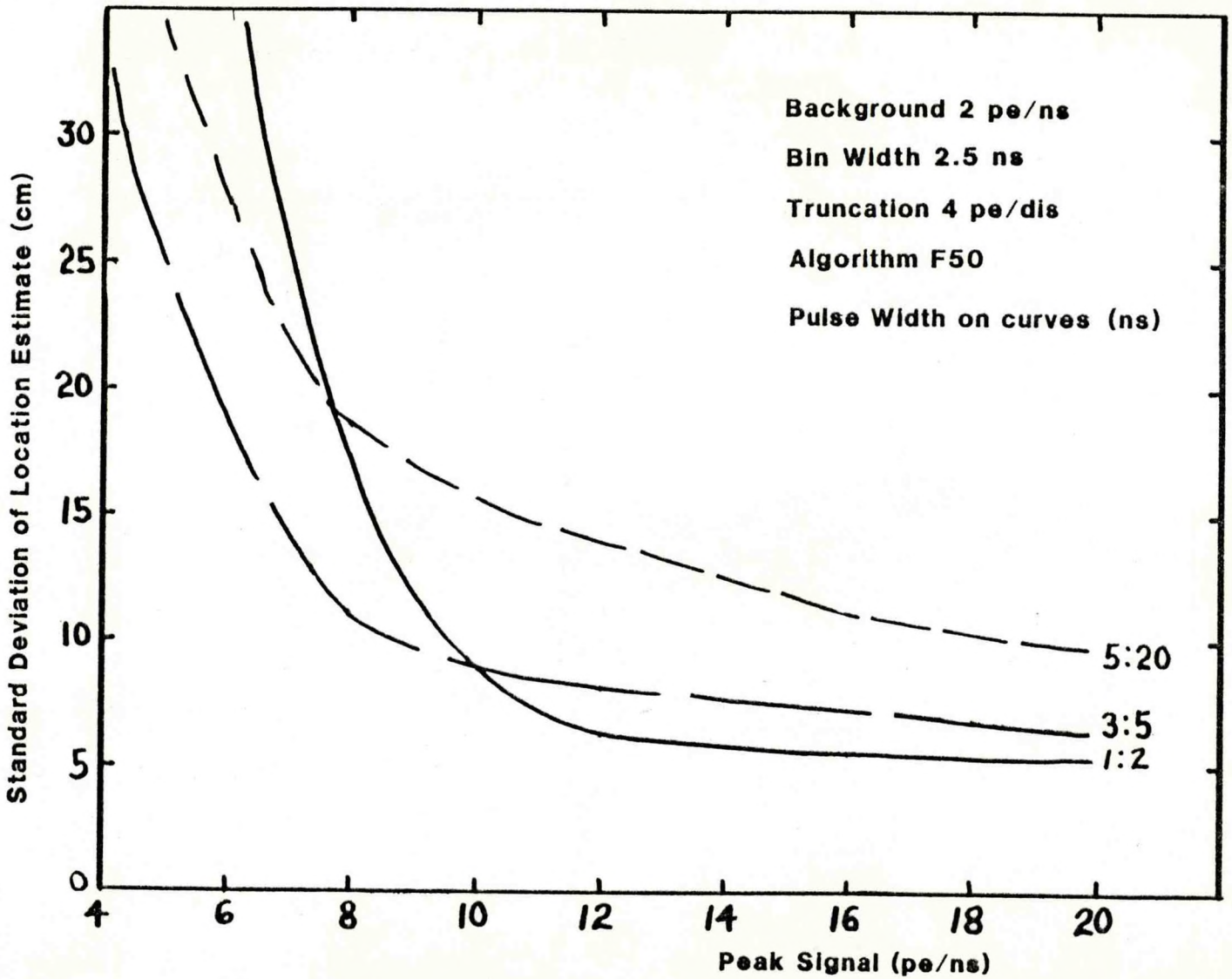

Figure 17. Pulse Width Effect for Nominal Bin

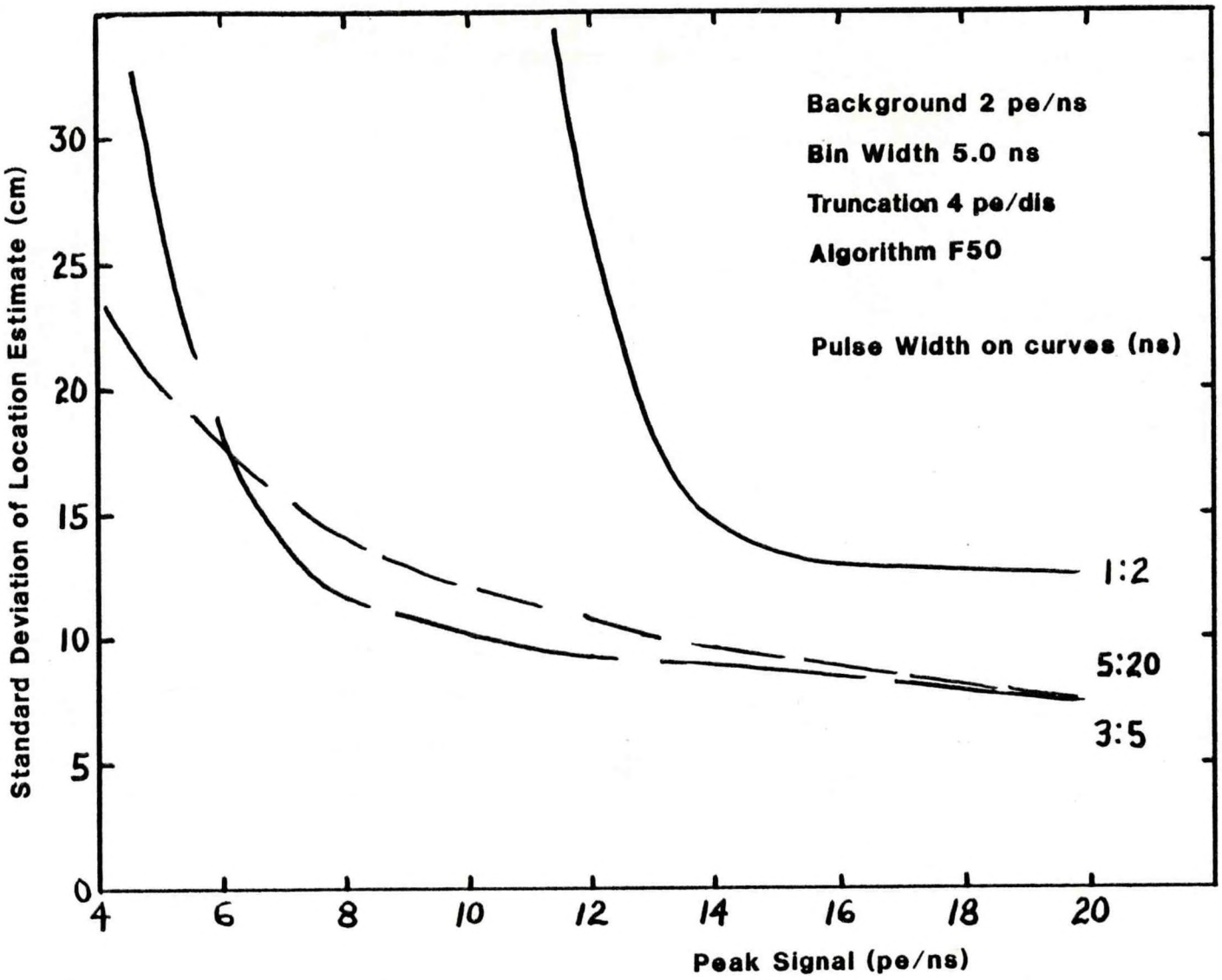

Figure 18. Pulse Width Effect for Wide Bin

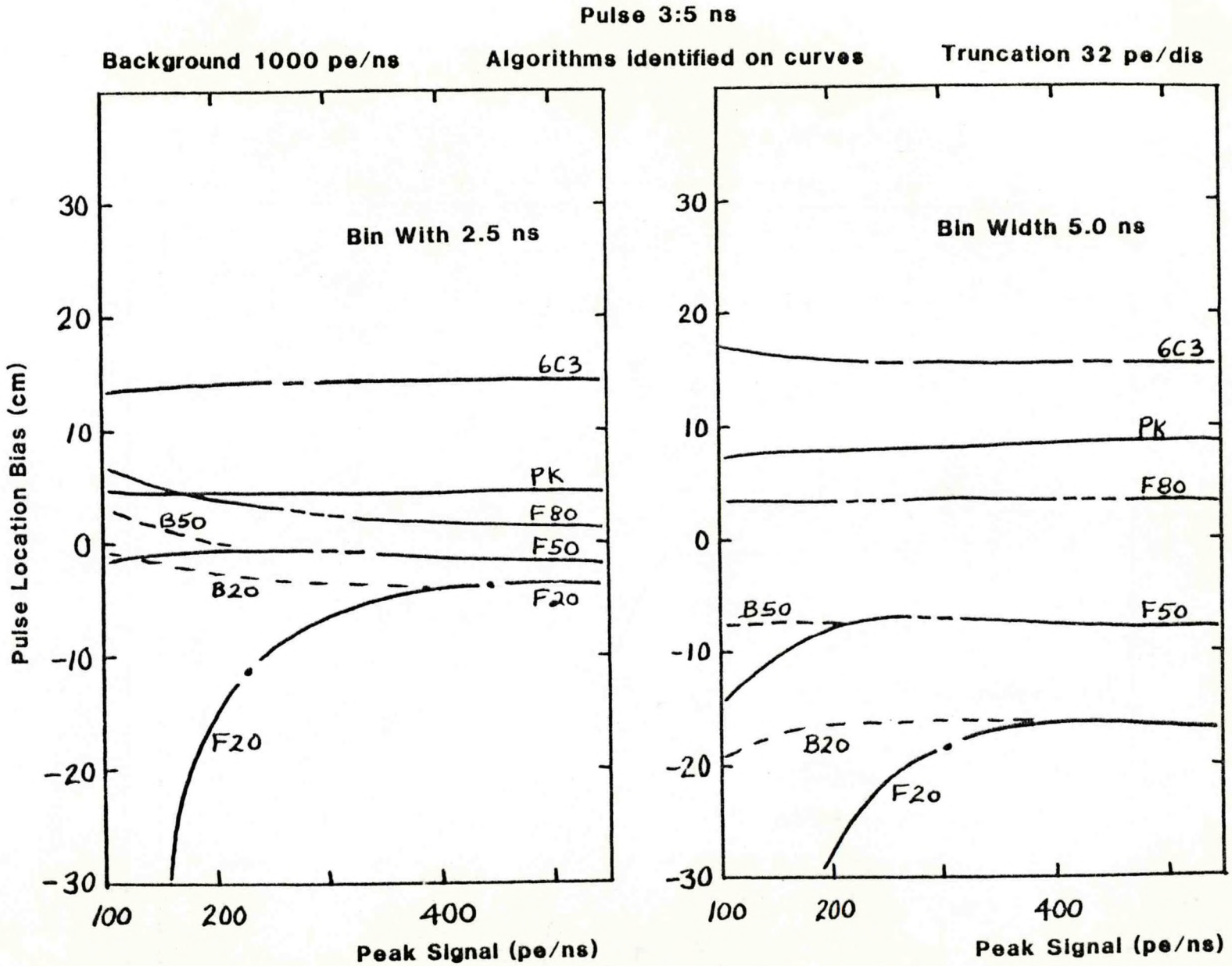

Figure 19. Biases for Nominal Pulse in Daylight

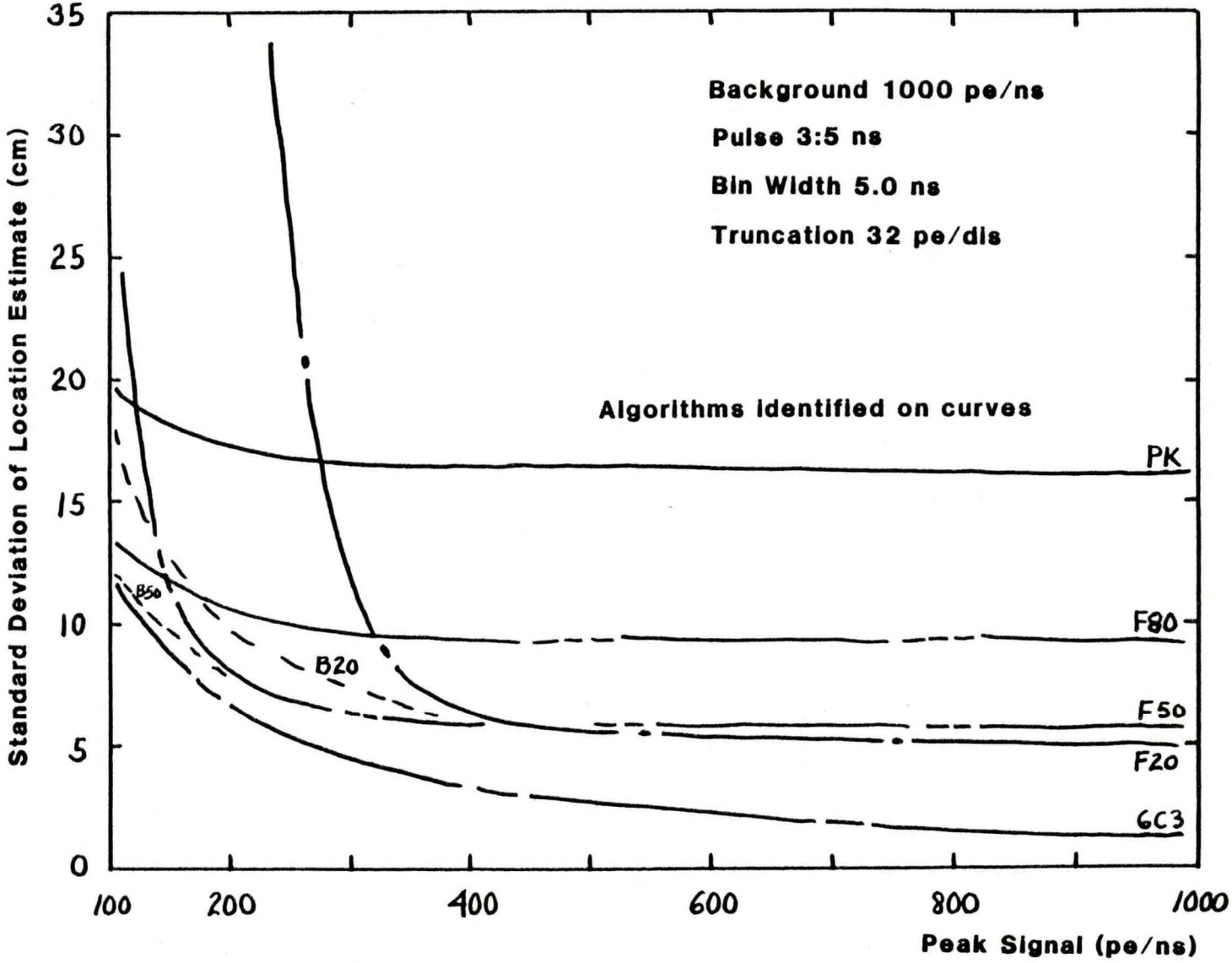

Figure 20. Precision for Nominal Pulse in Daylight

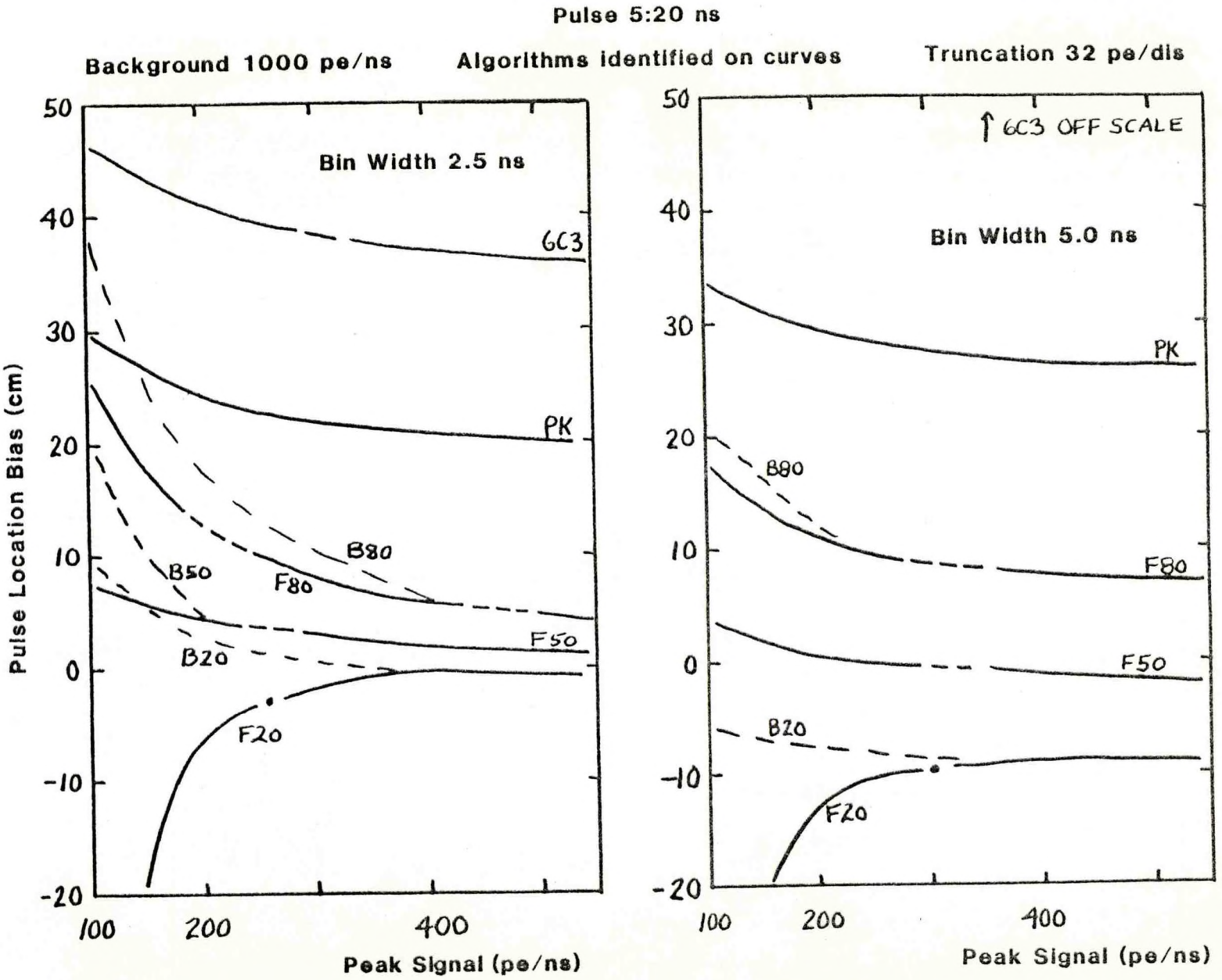

Figure 21. Biases for Stretched Pulse in Daylight

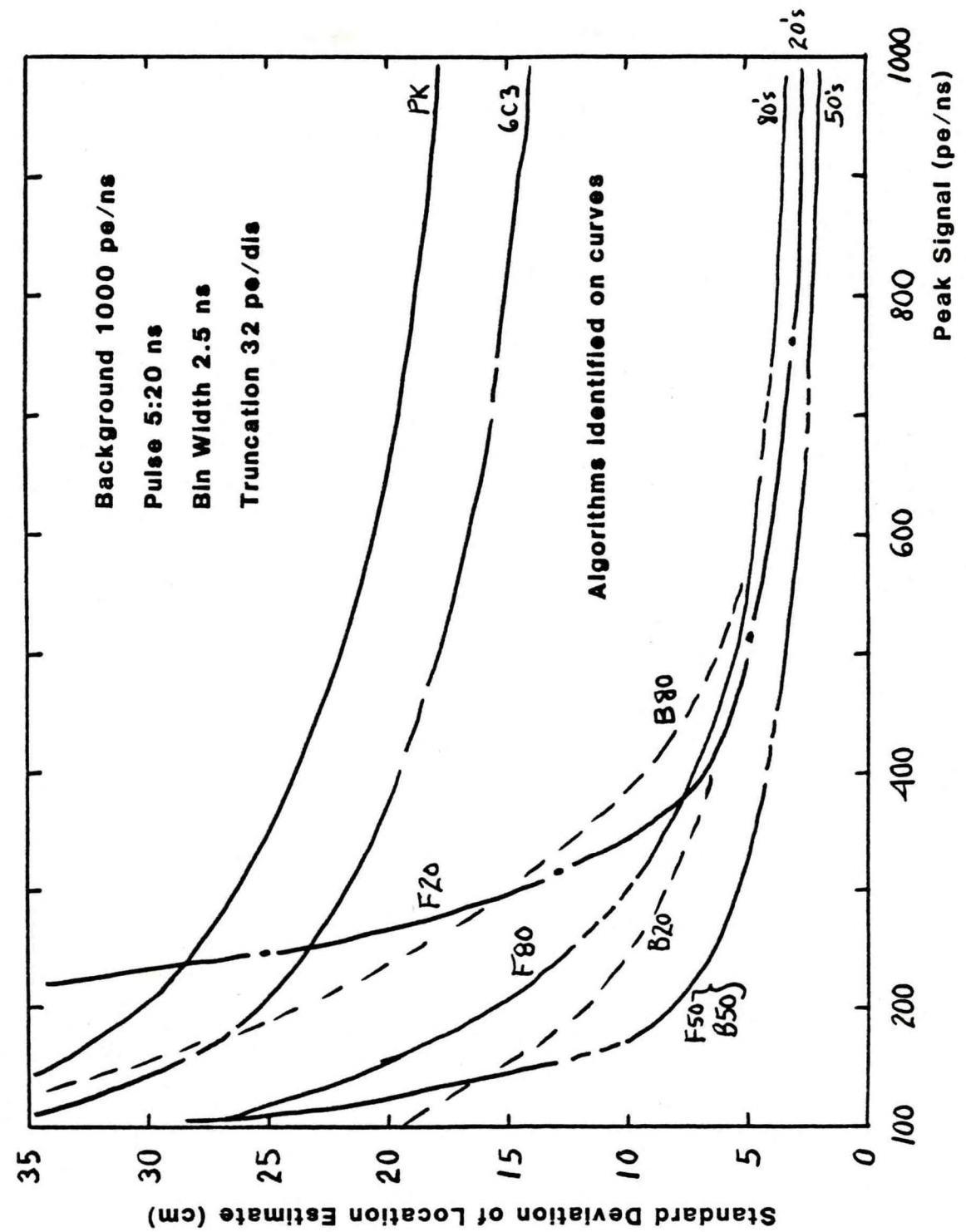

Figure 22. Precision for Stretched Pulse in Daylight

NOAA SCIENTIFIC AND TECHNICAL PUBLICATIONS

The National Oceanic and Atmospheric Administration was established as part of the Department of Commerce on October 3, 1970. The mission responsibilities of NOAA are to assess the socioeconomic impact of natural and technological changes in the environment and to monitor and predict the state of the solid Earth, the oceans and their living resources, the atmosphere, and the space environment of the Earth.

The major components of NOAA regularly produce various types of scientific and technical information in the following kinds of publications:

PROFESSIONAL PAPERS — Important definitive research results, major techniques, and special investigations.

CONTRACT AND GRANT REPORTS — Reports prepared by contractors or grantees under NOAA sponsorship.

ATLAS — Presentation of analyzed data generally in the form of maps showing distribution of rainfall, chemical and physical conditions of oceans and atmosphere, distribution of fishes and marine mammals, ionospheric conditions, etc.

TECHNICAL SERVICE PUBLICATIONS — Reports containing data, observations, instructions, etc. A partial listing includes data serials; prediction and outlook periodicals; technical manuals, training papers, planning reports, and information serials; and miscellaneous technical publications.

TECHNICAL REPORTS — Journal quality with extensive details, mathematical developments, or data listings.

TECHNICAL MEMORANDUMS — Reports of preliminary, partial, or negative research or technology results, interim instructions, and the like.

Information on availability of NOAA publications can be obtained from:

ENVIRONMENTAL SCIENCE INFORMATION CENTER (D822)
ENVIRONMENTAL DATA AND INFORMATION SERVICE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
U.S. DEPARTMENT OF COMMERCE

6009 Executive Boulevard
Rockville, MD 20852

NOAA--S/T 81-197