

Higher certainty of the laser-induced damage threshold test with a redistributing data treatment

Cite as: Rev. Sci. Instrum. **86**, 103106 (2015); <https://doi.org/10.1063/1.4932617>

Submitted: 03 July 2015 . Accepted: 26 September 2015 . Published Online: 13 October 2015

Lars Jensen, Marius Mrohs, Mark Gyamfi, Heinrich Mädebach, and Detlev Ristau



View Online



Export Citation



CrossMark

ARTICLES YOU MAY BE INTERESTED IN

[Revision of laser-induced damage threshold evaluation from damage probability data](#)

Review of Scientific Instruments **84**, 045108 (2013); <https://doi.org/10.1063/1.4801955>

[Wavelength dependence of femtosecond laser-induced damage threshold of optical materials](#)

Journal of Applied Physics **117**, 223103 (2015); <https://doi.org/10.1063/1.4922353>

[Determination of ultra-short laser induced damage threshold of \$\text{KH}_2\text{PO}_4\$ crystal: Numerical calculation and experimental verification](#)

AIP Advances **6**, 035221 (2016); <https://doi.org/10.1063/1.4945415>



JANIS

Janis Dilution Refrigerators & Helium-3 Cryostats
for Sub-Kelvin SPM

Click here for more info www.janis.com/UHV-ULT-SPM.aspx

Higher certainty of the laser-induced damage threshold test with a redistributing data treatment

Lars Jensen, Marius Mrohs, Mark Gyamfi, Heinrich Mädebach, and Detlev Ristau
Laser Components Department, Laser Zentrum Hannover e.V., Hannover 30419, Germany

(Received 3 July 2015; accepted 26 September 2015; published online 13 October 2015)

As a consequence of its statistical nature, the measurement of the laser-induced damage threshold holds always risks to over- or underestimate the real threshold value. As one of the established measurement procedures, the results of S-on-1 (and 1-on-1) tests outlined in the corresponding ISO standard 21 254 depend on the amount of data points and their distribution over the fluence scale. With the limited space on a test sample as well as the requirements on test site separation and beam sizes, the amount of data from one test is restricted. This paper reports on a way to treat damage test data in order to reduce the statistical error and therefore measurement uncertainty. Three simple assumptions allow for the assignment of one data point to multiple data bins and therefore virtually increase the available data base. © 2015 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution 3.0 Unported License. [<http://dx.doi.org/10.1063/1.4932617>]

I. INTRODUCTION

The measurement of the laser-induced damage threshold (LIDT) has a long history beginning shortly after the advent of the laser at the beginning of the 1960s. Over the years, a number of important lessons needed to be learned to derive a meaningful and reproducible damage threshold values. Among many other parameters, laser beam stability, beam size, and sample preparation require careful attention and a certain level of experience.¹

Most laser applications apply rather compact optical components. Also, a meaningful test of the power handling capabilities needs to be performed on an optic of identical quality. This includes substrate material, subsurface, and surface quality and finishes as well as coating design and deposition. As the substrate quality is highly dependent on the capability of cutting, grinding, and polishing of certain sizes, it is usually desirable to test samples of limited surface area. However, a statistical test calls for a maximum amount of test data to decrease uncertainty, which is contrary to the trend of compact high power lasers. In this paper, we report on a data treatment method to decrease error budgets of a LIDT test data evaluation without adding more data to the given set.

II. DATA REDUCTION

Concerning the general test procedure, the reader may consult the ISO standard 21 254 as the present paper builds on these procedures and data recording methods for the 1-on-1 and S-on-1 tests. All data presented (simulated and experimental) are produced with a continuous distribution on the fluence scale. The binning is always conducted after the sample has been irradiated with the whole data set and not already during the testing phase by retesting at certain fluence levels with pre-selected intervals. In the experiment, the motivation for this protocol is to be flexible in distributing the remaining test sites over the fluence scale, particularly for samples of unknown threshold level.

The following three basic assumptions comprise the foundation for the presented data reduction method.

1. An undamaged test site would have also survived when irradiated at lower fluence.
2. A damaged test site would have also been damaged when irradiated at higher fluence.
3. In one of these virtual additional tests, the given defects were identically distributed within the Gaussian beam profile as in the actual tested site. (The number of assumptions reduces to the first two when testing with a top hat beam profile or when the test beam size is sufficient to cover the given defect ensemble within the test area.)

With these three assumptions, it is a direct consequence that the test data of a certain fluence interval or data bin can also be assigned to further bins in order to significantly increase the total amount of data relevant for the subsequent data reduction procedure. This approach is graphically shown in Fig. 1. In this view graph, a data set is shown which has been taken in a test with continuous laser fluence values. After collection, the data were arranged into equally sized fluence bins, whose width is chosen with respect to the intervals on the fluence scale and not concerning the count of data points per bin. From these bins, the damage probability p_i is calculated as the ratio between the amount of damaged sites n_d and total number of test sites in the respective bin (damaged and non-damaged n_{nd}),

$$p_i = \frac{n_d}{n_d + n_{nd}}. \quad (1)$$

In this example, only the damage probabilities in the transition range of the full data set are marked with the vertical dashed lines. All other bins will give probabilities equal to 1 or 0, as only one of the two possible states is included in the respective bin. For the sake of simplicity, these bins are not shown in Fig. 1.

With this procedure, the average number of data points per bin (sample diameter $d = 25$ mm) is increased from about 10 to roughly 40. Assuming a statistical error of $1/N$, the reciprocal amount of data points per bin, the statistical probability error

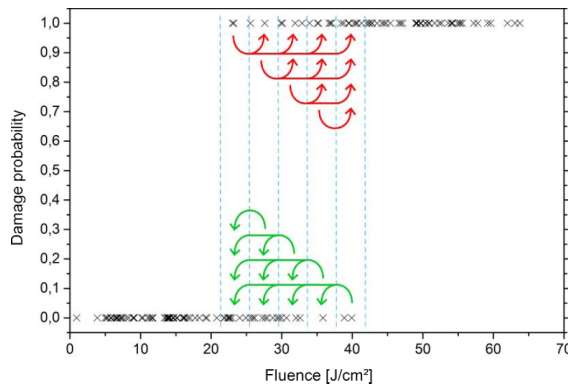


FIG. 1. Test data with a continuous distribution on the fluence scale assigned not only to its original data bin. Data of damaged test sites are also assigned to bin of higher fluence and data of survived test sites are also assigned to bins of lower fluence.

is reduced from $\sim 10\%$ to $\sim 2.5\%$.² And just as suggested by Hildenbrand *et al.*, the error bars are limited to probability values between 0 and 1, as it is not a physical assumption to allow values of $p_i > 1$ and $p_i < 0$.³ The result of this data reduction is shown as a before and after comparison in Fig. 2. The fitting routine to derive the damage threshold value is addressed in Secs. III and IV.

In total, this procedure is based on the experimental settings used in a standard damage threshold test according to the ISO standard. In a test on a sample of 25 mm in diameter irradiated under an angle of 0° , this will result in roughly 160 test sites or data points for the evaluation. This number is chosen with respect to a certain distance between test sites to avoid cross talk and a certain distance from the sample edge to exclude edge effects in the data. Evaluating data sets of this size will give the stated improvements in error budgets.

Also, two additional advantages of this method are the monotonous increase of the damage probability and the possibility of more bins without increasing the statistical error substantially. Since all data points of a certain fluence interval and damage status can be assigned to neighboring interval as well, the derived damage probability will result in a monotonous characteristic of the probability distribution. Taken most experimental sets of data, this is not the case when evaluating each interval as a completely independent

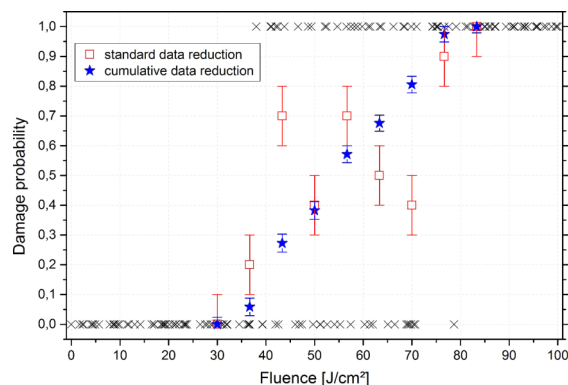


FIG. 2. Comparative representation of the established and the cumulative evaluation of the damage probability based on an ISO based test protocol. In this case, fitting of the damage probability points is much easier, if the probabilities were calculated with the new method.

bin. Particularly of interest for a non-linear regression of the probability data, a rather dense data distribution is preferable in order to find the correct model and fit parameters to describe the given optical component. However, increasing the number of fluence intervals will directly increase the corresponding error because the amount of data per bin will decrease. This method allows for a higher flexibility when evaluating the given data without a significantly increased error.

One characteristic of this method is that the damage probability of the first fluence bin with a non-zero damage probability will always get a lower damage probability than with the standard method, because it can only gain additional non-damaged sites, but no damaged ones. The opposite case applies for the last bin with damage probability below 1 which can only gain damaged sites. Damaged test spots at relatively low fluences and non-damaged test spots at relatively high fluences have a higher impact on the results and therefore a higher weighting. This might lead to a steeper gradient of the damage probability and thus to slightly higher damage threshold, especially if the sample exhibits a slow increase in damage probability at lower fluences and a fast increase at higher fluences. Within this issue, it is recommended to apply this redistributed data treatment only to complete data sets, with recorded data below the damage threshold, within the transition range, and also beyond the 100% damage probability.

III. CASE STUDY

To illustrate the benefits of the suggested approach, four virtual, randomly generated data⁹ sets have been analyzed concerning the impact of this method on the fitting errors and correlation to the data. Although the damage threshold in these four respective data sets is very similar, the distribution of the data is different for each case, just as it is the case in real experiments. Again, the data are distributed randomly over the fluence scale and not assembled in distinct intervals. The data are assigned to the fluence bins as last step before the damage probability is calculated for each bin. These sets of data have been evaluated with the standard procedure and with the presented cumulative data reduction. As a second step, each of these two probability distributions is fitted with a non-linear regression. Out of a number of possible models⁴⁻⁷ (just to name a few), the work by Porteus and Seitel⁴ has been chosen to be the basis for fitting both versions of the damage probability distribution. The result is illustrated in Fig. 3 for each data set. In these view graphs, for the sake of clarity, only the fit function of the cumulative reduced probability distribution is shown, and the complete test results are listed in Table I.

Although the damage threshold in these four respective data sets is very similar, the distribution of the data is different for each case, just as it is the case in real experiments. For each of these four examples, it is obvious that the cumulative data reduction introduces a monotonous distribution of damage probability. Usually, at least constant but also reversed trends are found in each evaluation of laser damage test data. Now, monotonous trends with significantly reduced error bars (based on the amount of data points in each bin) lead to a significantly increased correlation between fit function and probability distribution. This higher correlation is directly linked to

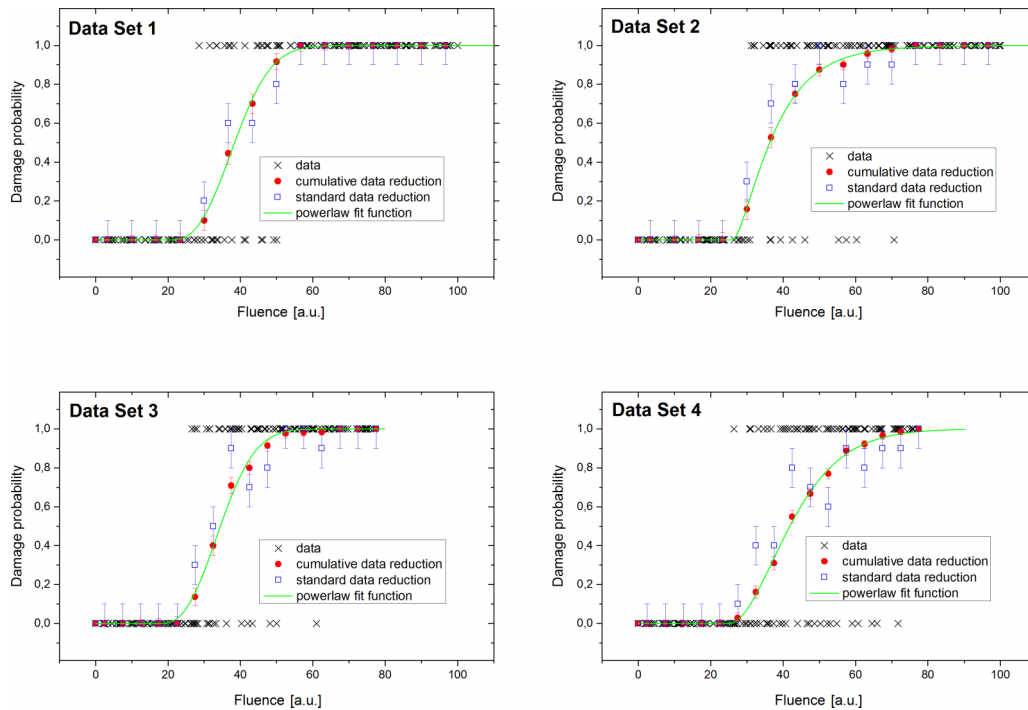


FIG. 3. Damage threshold analysis with the standard and cumulative data reduction of four randomly simulated data sets. For the sake of clarity, only the fit function of the cumulative reduced probability distribution is shown.

a higher certainty of the test results. To assess these results, in Table I, the fitting errors and the standard deviation between probability data and fit function are listed. The fitting error of the LIDT is given by the standard deviation of the fit coefficients in the Levenberg-Marquardt algorithm.^{8,10} In this case, it is usually reduced by a factor of 2 or 3 and the standard deviation between data ensemble and fit function is approximately five times lower when using the new cumulative method.

Since the laser-induced damage threshold is the physical characteristic of interest in such a test, the impact of this progress in error development on the LIDT is the final measure for improvement. The error of the LIDT value was reduced by more than 50%, and statistically, the fitting error of the LIDT of well below 3% is achieved.

IV. REPRESENTATION OF THE TESTED DAMAGE PROBABILITY

The main result of the previous section is that a lower standard deviation can be achieved when using the new method for the calculation of damage probabilities and fitting them with the powerlaw function. This is usually desirable, but

it does not necessarily mean that the resulting probability distribution is a better representation of reality than in the case of the standard method. It remains the question which method leads to the better representation.

To test this, another ten sets of data were generated and the standard deviation for the resulting probability distributions in comparison to the powerlaw function was calculated, which is shown in Fig. 4. The probability function used for the generation of the test spots (dashed green line) represents the “reality” and by randomly generating damaged or non-damaged test sites from it, a measurement of this “reality” was simulated. To achieve a broad spectrum of cases, the defect density, the threshold, the number of generated sites, and the factor p of the powerlaw function were varied. The standard deviation was calculated for both methods with five and ten fluence intervals, respectively, by comparing the damage probabilities calculated from the generated virtual test spots and the model curve used for generation of data. The results can be found in Table II.

The results in Table II clearly show that the new cumulative method of calculation leads to a better representation of reality than the established method. There was not a single

TABLE I. Evaluation errors and standard deviations of the four presented data sets with the standard and cumulative reduction technique.

	Cumulative LIDT error (%)	Standard LIDT error (%)	Cumulative standard deviation	Standard deviation
Data set 1 (Fig. 3)	1.3	3.9	0.0138	0.0507
Data set 2 (Fig. 3)	2.6	5.0	0.0108	0.0620
Data set 3 (Fig. 3)	1.4	4.7	0.0137	0.0636
Data set 4 (Fig. 3)	3.6	10.0	0.0128	0.0746
Averaged over all data sets	2.2	5.9	0.0128	0.0627

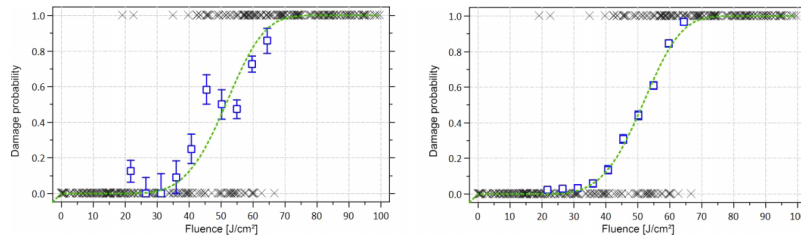


FIG. 4. Comparison of damage probabilities deduced from randomly generated damage test data based on a given distribution function—left: standard data reduction—right: new cumulative data reduction. The depicted data correspond to data set 5 in Table II. The visible green curve shows the fixed model curve that was used to generate the virtual test spots visible at the top and the bottom.

case where the standard method delivered a better determined probability distribution. The calculated standard deviation was in average approximately five times lower when using the new method. Although our tests were limited to only ten data sets, we can still conclude that the cumulative technique is advantageous when trying to find the respective probability distribution.

V. DIVERSE DEFECT ENSEMBLES

As already discussed in Sec. IV, the described cumulative data reduction technique does not broaden the damaged probability distribution although the data are used in multiple fluence bins. This is the case because only survived or damaged test sites are counted in the lower or higher bins, respectively. Never are test sites of different status assigned to a bin in which they have not been tested.

A second part of laser damage phenomena, where this is a helpful fact, is the evaluation for more than one defect type per defect ensemble. As stated a few times before,^{5,6} if a defect ensemble consists of more than one class of defects, the damage probability function will show a discontinuity and the different parts of the distribution have to be fitted with different parameters revealing diverse defect densities. Given that a sufficient number of fluence bins has been chosen to resolve these discontinuities, the cumulative data reduction keeps them in position (on the fluence scale). Usually, the presence of an additional defect class will be revealed more clearly because of the monotonous nature of the damage probability when derived with this method. One example for this is presented

in Fig. 5. Two different numbers of fluence bins have been selected to evaluate a given data set (Fig. 5(a): 20 bins and Fig. 5(b): 35 bins). The set consists of 500 test positions, generated randomly in the same procedure as discussed in Sec. III. The data reduction to derive the damage probability was conducted using the standard procedure as well as the cumulative approach. When selecting an insufficient number of bins (Fig. 5(a)), the discontinuity is not clearly resolved and can only be estimated. When applying the standard evaluation, the distribution of the damage probability for each bin is spread more widely, and the two separate parts cannot be resolved. The cumulative approach provides the possibility to increase the number of bins within the evaluation without significantly increasing the statistical error of the damage probability. This is shown in Fig. 5(b). Although this example already has access to a data set of 500 test sites, the amount of 35 bins introduces error bars of up to ± 10 percent in the standard evaluation. This directly results again in an indistinct probability characteristic which does not clearly show the discontinuity of a defect ensemble with two classes. The cumulative algorithm, however, provides data with small error bars and reveals a clear transition from one defect class to the other. Both parts of the distribution can be fitted separately, and defect densities are accessible for further optimization of the thin film or optical material.

VI. SHIFTED DATA BINS

With continuously distributed data on the fluence scale, the division in a certain number of bins is fully determined by

TABLE II. Standard deviations for both methods when comparing the function used for generating data and the resulting probability distributions.

Data set	Standard method (5 intervals)	Cumulative method (5 intervals)	Standard method (10 intervals)	Cumulative method (10 intervals)
1	0.3176	0.1722	0.4819	0.2362
2	0.3658	0.0389	0.5169	0.3448
3	0.4932	0.0258	1.0868	0.0658
4	0.5283	0.0591	1.0852	0.0597
5	0.2212	0.0376	0.4142	0.0814
6	0.2913	0.0757	0.5880	0.0858
7	0.3215	0.0269	0.9117	0.0874
8	0.3637	0.0599	0.6904	0.0737
9	0.3387	0.1361	0.6445	0.1617
10	0.1223	0.0275	0.5101	0.1159
Average	0.3364	0.0660	0.6930	0.1313

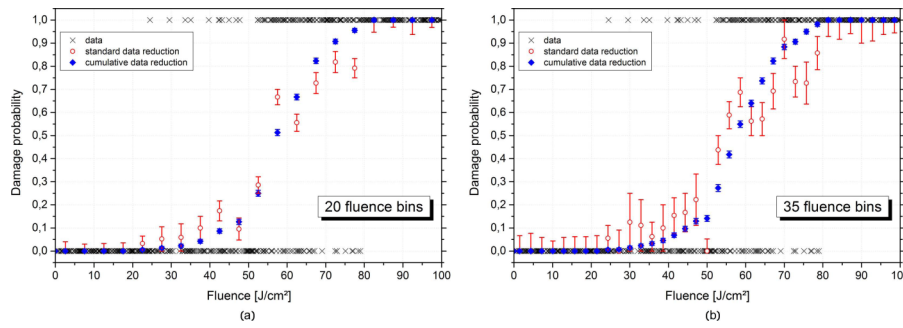


FIG. 5. Damage probability curve describing a defect ensemble of two defect classes—evaluated with 20 and 35 fluence bins.

the evaluation or the specific algorithm. This combined with the error bar extended over the full width of the bin adds high uncertainty. By continuously shifting the boundaries of these bins over the data set, it is made visible how this affects a possible shift of damage probability on the fluence scale when randomly dividing the data set into bins. This procedure is illustrated for a damage probability curve in Fig. 6. In the table on the left, all data in the transition range are listed—meaning the fluence range in which damaged and survived sites have been observed simultaneously. This particular full data set has been divided into 7 bins of energy density, and of course, the corresponding boundaries can be chosen with a certain degree of freedom. In the mentioned shift procedure, the bins have been assigned a width which includes 10 recorded fluence values and these bins have been shifted through the full data set resulting in 6 bin sets of 7 bins (5 bins plus one additional bin with a damage probability of 0 and one with a damage probability of 1) each. The damage probability of each bin was then evaluated according to the cumulative data reduction. It is obvious that the probability transition from 0 to 1 of this data set is not broadened by taking and treating the test data with this approach. Of course, this is only shown in this example for this publication. Since the whole algorithm is a numerical one for real experimental data, it is not possible to analytically show the impact of the procedure as generally valid.

However, working with this routine and comparing it for numerous data sets showed repeatedly the same result. To also express this in numbers, in Table III, the fitted LIDT values for each of the shifted bin sets are listed. The before mentioned powerlaw model has been used to fit these sets using a constant p factor (see Ref. 4).

At the bottom of the table, the average values of the fitted LIDT and the fit errors are stated. Additionally, the standard deviation of the six LIDT values of the respective bin groups is calculated, and it is significantly smaller than the fitting error average. Usually, it is also smaller than the fit errors of each single data set. This example illustrates the experience with this approach on laser damage data treatment and its error development.

Since this paper reports on the data reduction technique, the error budget in the discussed examples for LIDT tests will only consider the statistical uncertainty. For the experiment, the operator always has to include the fluence fluctuations for the specifically applied laser as well. Depending on the stability of the laser, this can add a substantial part of the overall error. Total absolute errors of LIDT measurements are usually in the range of 10%-20%.

One possible approach to derive a damage threshold without applying a fit function is to utilize this bin shifting procedure on the set of test data including the cumulative data reduction over the respective bins. The bin assigned to the highest fluence showing a damage probability of 0 can be considered as significant for the damage threshold, and the error to be applied to this result would be the width of the bin on the fluence scale. However, this approach will only give the value for the 0% damage probability and no insights into defect distributions or higher percentages. Also, this procedure will not account for the fact that many samples never show a damage probability of 0, just very close to 0. However, it would take out the uncertainties introduced by the choice of the model and of the start parameters for the fitting routine. Only the beam parameters and the bin width on the fluence scale would determine the total error budget.

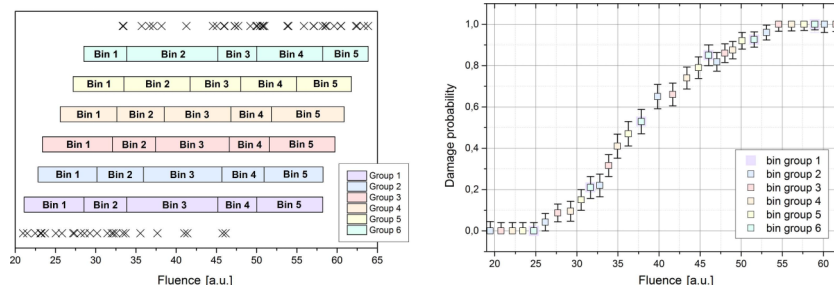


FIG. 6. Left picture: schematic representation of the bin shifting procedure. Right picture: continually shifted bins of 1-on-1 damage data on the fluence scale. The bins were shifted by 2 test spots (crosses). With this procedure, identical damage probability distributions are derived to show that continuously distributed data on the fluence scale does not increase the fluence error.

TABLE III. Fitted LIDT including error and standard deviation of continually shifted data bins as shown in Fig. 6.

	Fitted LIDT (a.u.)	Fit error (%)
Bin group 1	25.2	1.52
Bin group 2	25.7	3.92
Bin group 3	24.6	4.12
Bin group 4	24.9	3.04
Bin group 5	25.3	2.10
Bin group 6	25.4	1.94
Average	25.2	2.77
Standard deviation	1.54%	

VII. CONCLUSION

This publication reports on a data reduction technique in laser-induced damage testing which offers a reduced error budget and therefore lower uncertainty in the evaluation of laser damage data. It has been shown that a data set can be virtually increased by assigning single data points to more than one data bin based on 3 simple physical assumptions. Based on these assumptions, the available amount of data is roughly quadrupled and also a monotonous trend of the damage probability is introduced to the curve. Fitting correlations are enhanced and damage threshold determination shows now lower uncertainty. Additionally, separate defect classes of one defect ensemble are now accessible also with a data set of reasonable size. In summary, the statistical error of the LIDT measurement is clearly reduced by applying the cumulative data reduction technique. The reduction of the error can be averaged to a factor of 2–3. Of course, the error budget depending on the stability of the laser source and beam size and shape is still not affected by this algorithm.

ACKNOWLEDGMENTS

The authors would like to acknowledge Stefan Schrameyer and Jens Vogel for their contribution to this field. In addition, the support of the German Ministry of Education and Research (BMBF) is acknowledged within the frame of the project Ultra-LIFE under Contract No. FKZ:13N11558.

¹ISO 21254:2011 “Test methods for laser-induced damage threshold,” International Organization for Standardization.

²S. Schrameyer, M. Jupé, L. Jensen, and D. Ristau, “Algorithm for cumulative damage probability calculations in S-on-1 laser damage testing,” *Proc. SPIE* **8885**, 88851J (2013).

³A. Hildenbrand, F. R. Wagner, H. Akhouayri, J. Y. Natoli, and M. Commandré, “Accurate metrology for laser damage measurements in nonlinear crystals,” *Opt. Eng.* **47**(8), 083603 (2008).

⁴J. Porteus and S. Seitel, “Absolute onset of optical surface damage using distributed defect ensembles,” *Appl. Opt.* **23**, 3796–3805 (1984).

⁵J.-Y. Natoli, L. Gallais, H. Akhouayri, and C. Amra, “Laser-induced damage of materials in bulk, thin-film and liquid forms,” *Appl. Opt.* **41**, 3156–3166 (2002).

⁶H. Krol, L. Gallais, C. Crèzes-Besset, J.-Y. Natoli, and M. Commandré, “Investigation of nanoprecursors threshold distribution in laser damage testing,” *Opt. Commun.* **256**, 184–189 (2005).

⁷L. Lamaignère, S. Bouillet, R. Courchinoux, T. Donval, M. Josse, J.-C. Ponetta, and H. Bercegol, “An accurate repeatable and well characterized measurement of laser damage density of optical materials,” *Rev. Sci. Instrum.* **78**, 103105 (2007).

⁸J. Alper and R. Gelb, “Standard errors and confidence intervals in nonlinear regression: Comparison of Monte Carlo and parametric statistics,” *J. Phys. Chem.* **94**, 4747–4751 (1990).

⁹These data sets were generated with random numbers between 0 and 1. If the random number was above the previously chosen model curve, the virtual test spot was set to “not damaged” and for random numbers below the model curve to “damaged.” This leads to realistic distributions of virtual test spots, which fit the model curve very well for high numbers of test spots.

¹⁰In particular the diagonal elements of the covariance matrix state the square root of the standard deviation for each fitting coefficient.⁸ In this case, the damage threshold and the defect density are the coefficients relevant for this fit (in addition to the p-factor).⁴