SAND2017-4831R

Sandia National Laboratories

Operated for the U.S. Department of Energy by
Sandia Corporation

Albuquerque, New Mexico 87185

date: May 2, 2017

to: Joshua Stein

from: Clifford Hansen, Sandia National Laboratories
Dirk Jordan, National Renewable Energy Laboratory

subject: Sample size for PV lifetime project

The PV lifetime project must select a sample size (number of PV modules) for each PV system to be deployed. In this memorandum, we show how the uncertainty in measured degradation depends on the selected sample size.

Assumptions

The PV lifetime project will deploy several samples of modules in an experiment to determine degradation over time. We want to estimate the sample size number of modules needed to statistically determine the degradation rate. We make the following assumptions:

1. The sample of modules represents a random sample from one Pmp bin so that variation in Pmp among modules can be viewed as random.
2. Module degradation rates are not correlated with Pmp.
3. Random error in Pmp measurement is small compared to variation in Pmp among modules.
4. Systematic error in Pmp measurement (accuracy) does not change, or can be controlled by calibration to a constant reference.

Pmp is measured for each module in the sample at time zero. Pmp is measured in the same conditions for each module in the sample after T years. The difference in Pmp divided by the time is the observed degradation rate:

$$r_i = \frac{P_{T,i} - P_{0,i}}{T} \quad (1)$$

For comparison among modules in the sample, it is convenient to convert degradation rate from W/yr to %/yr by dividing each module's degradation rate by its nominal power rating. The above assumptions are made in order that $r_i, i = 1, \ldots, N$ can be considered as a random sample of degradation rates from the distribution of a random variable *R*, representing the population of degradation rates, with unknown mean μ and variance σ^2.

Confidence intervals for the measured degradation rate

We can use published research to obtain a conservative estimate of the standard deviation σ of *R*. A review of reported degradation rates [1] shows that degradation rates pooled over many module types, lengths of deployment, and manufacturers form a right-skewed distribution with a long tail. No standard deviation is reported in [1]; however, we can estimate standard deviation by adopting a log-normal model for the data. For the compilation of rates illustrated in Figure 2a, with mean of 0.8% and median of 0.5%, we obtain a standard deviation of 1.0% for the lognormal model of the data in [1].[1] We view 1.0% as a conservatively large estimate of the standard deviation in degradation rate for a sample of modules of comparable vintage and technology from a single manufacturer. It is reasonable to assume that a manufacturer's process controls result in much less randomness than is observed in the data in [1].

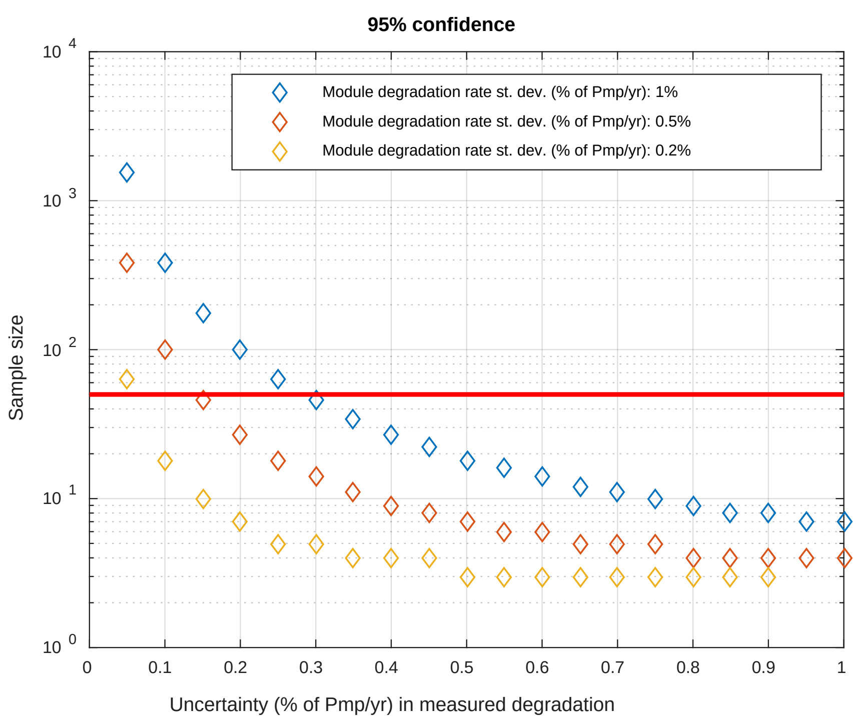
Using this conservative estimate of the standard deviation σ we can estimate the sample size needed to determine the degradation rate with a certain level of precision. The level of precision does not depend on the actual value of the underlying population degradation rate. The sample size needed to obtain a $100(1-\alpha)\%$ confidence interval about the sample mean with expected width $2h$ is given by

$$N = \left(\frac{\sigma t_{\alpha/2,N-1}}{h} \right)^2 \qquad (2)$$

where $t_{\alpha/2,N-1}$ is the $1-\alpha/2$ percentile of the t-distribution with $N-1$ degrees of freedom.

Guidance for selecting sample size

In Eq. 2, the terms σ and h have the same units. It is convenient to express these terms in units of % of Pmp per year. For a confidence of 95% ($\alpha = 0.05$), solving Eq. 2 for N results in samples sizes shown in the figure below.



---

[1] The lognormal model has parameters mu and sigma. The data have median = 0.5% and mean = 0.8%. mu = ln(0.5), sigma = sqrt(2×ln(0.8/0.5)). Variance is (exp(sigma^2) – 1)×(exp(2×mu + sigma^2)); standard deviation is sqrt(variance).

For example, suppose the project selects a sample of size 50 (indicated by the red line on the figure). If the degradation rates for the underlying module population vary significantly, with a variance $\sigma^2 = 1.0\%$, then the 95% confidence interval around the sample mean degradation rate is 0.3% (% of Pmp/yr). In this case, if the sample degradation rate is measured as 0.45%/yr, then the estimated population degradation rate is between 0.15% and 0.75% with 95% confidence. This interval is conservatively wide, deriving from the assumption of $\sigma^2 = 1.0\%$. After the sample of degradation rates is available, the sample variance may be used instead of $\sigma^2 = 1.0\%$ to estimate the population variance, and the estimated population variance can be used in turn to estimate an uncertainty interval which is likely to be much narrower.

1. Jordan, D.C. and S.R. Kurtz, *Photovoltaic Degradation Rates—an Analytical Review*. Progress in Photovoltaics: Research and Applications, 2013. **21**(1): p. 12-29.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.