SAND2022-14925C

A Statistical Assessment of Zener Diode Behavior Using Functional Data Analysis

Xiaoxia Champon[1,2], David Angeles[3], Thomas Buchheit[1], David Canfield[1], J. Derek Tucker[1], Jason Adams[1]

[1]Sandia National Laboratories, USA, [2]North Carolina State University, USA, [3]Ohio State University, USA

Abstract

This paper presents an assessment of electrical device measurements using functional data analysis (FDA) on a test case of Zener diode devices. We employ three techniques from FDA to quantify the variability in device behavior, primarily due to production lot and demonstrate that this has a significant effect in our data set. We also argue for the expanded use of FDA methods in providing principled, quantitative analysis of electrical device data.
(Keywords: Electrical Device Modeling, Functional Data Analysis)

Introduction

The principled statistical analysis of electrical device data can yield important insights and influence decisions in numerous areas related to device technology and manufacturing. This paper describes the statistical analysis of a set of electrical measurements obtained from 193 MMSZ522BT1G Zener Diodes. These diodes span two production lot date codes, and electrical measurements which characterize the operational behavior of these parts are current-voltage (I-V) sweeps. Fig. 1 displays the I-V curves from both production lots. From a visual perspective, the data presents with an unremarkable, relatively uniform distribution of results.

The primary goals of our analysis are to investigate whether device behavior differs across production lots and to tease out any other useful process related statistical characteristics from this data set. This paper employs techniques from functional data analysis (FDA) to provide a formal assessment of the behavior of these devices.

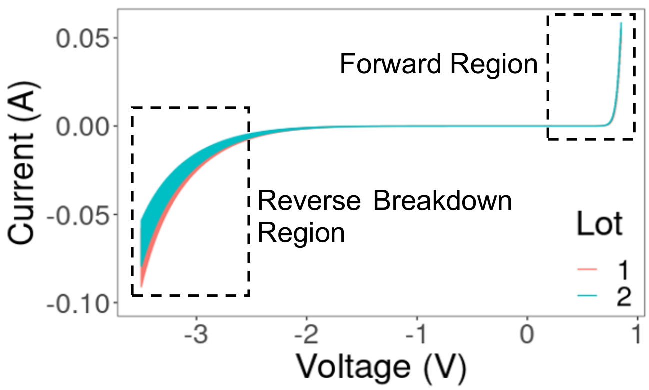

Figure 1: I-V curves from Zener diodes colored by production lot

Analysis of Electrical Device Data

A. Functional Data Analysis

Functional data vary continuously across an independent variable. Because our data set of I-V sweeps are generated by forcing voltage and measuring the current, it is natural to treat these curves as functional observations where current is considered as a function of voltage. We can express our data as $y_{ij}=f_i(t_{ij})$, where i indexes the subjects and j indexes the independent variable. In this paper, y represents the current, t represents the voltage, and the subjects correspond to the individual measured devices.

B. Amplitude and Phase Distances

A key aspect of any analysis of functional data is the characterization of two sources of variability. These are *amplitude*, or y-axis, and *phase*, or x-axis, variability. Ref. [1] details a method of separating amplitude and phase variability and describes two functional distance metrics – amplitude distance and phase distance. These distances allow us to understand how different two functions are from one another in terms of amplitude and phase variability.

As an initial exploratory step in our analysis, we compute the pairwise amplitude and phase distances for the I-V curves. When visualized, the pairwise distance matrices look similar. For this reason, we show only the amplitude distances in Fig. 2. The top left block of the matrix displays the distances between devices in production lot 1 while the bottom right block contains the distances for devices in production lot 2. The off-diagonal blocks then contain the between-lot comparisons (note that these blocks are transposed versions of each other).

If the production lots were producing identical devices, we would expect a nearly uniform matrix. However, the bands of higher pairwise distances observed in Fig. 2 indicate that there is a subset of devices in production lot 1 whose I-V curves differ substantially from the rest of the devices. This is an informative result suggesting that more analysis should be pursued to better understand what is driving the observed differences.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & 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.

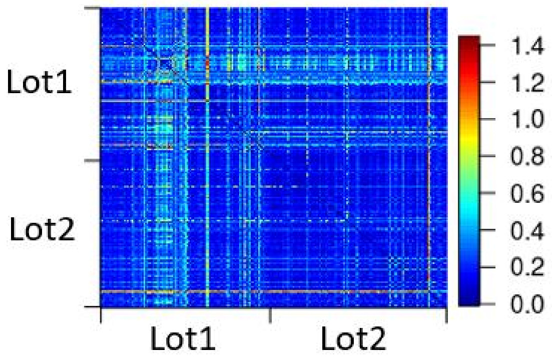

Figure 2: Amplitude pairwise distance

C. Functional Principal Components Analysis

In standard analysis of multivariate data, principal components analysis (PCA) is a ubiquitous tool for understanding the variability in the data and reducing dimensionality [2]. These same goals are pursued for functional data using functional principal components analysis (FPCA). Implementation details of FPCA can be found in chapter 8 of [3].

FPCA on the Zener diode measurements revealed that the first functional principal component (FPC) captures more than 99% of the variance in the data. The first two FPCs, capture over 99.99% of the variance. In Fig. 3, where the first two PCs are shown, the interpretation of these components is clear. The first component, denoted FPC1, describes the variability of the devices in the reverse breakdown region, identified in figure 1. FPC2 describes variability in the forward region. While this interpretation may not seem to provide much more information than a visual inspection of the data, we can learn a lot more when reducing the dimensionality of the data.

By computing the inner product of a centered I-V curve with an FPC, we obtain the corresponding FPC score. Fig. 4 plots FPC1 and FPC2 scores for all I-V curves. Two linearly separable clusters emerge that reveal a clear distinction in the behavior of

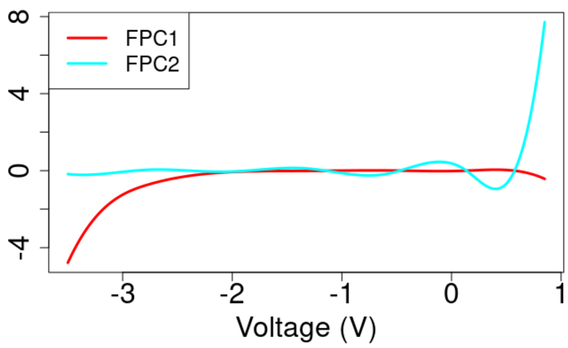

Figure 3: Functional principal components

devices between the two production lots (note that our implementation of FPCA is production lot agnostic). Interpreting the FPCs, it appears that relative to production lot 1, devices from production lot 2 tend to have smaller absolute (i.e., less negative) current values in the reverse breakdown region while having larger current values in the forward region.

Clearly, this implementation of FPCA sheds light on the specific device behavior influencing the difference we have observed between the production lots. In general, FPCA is a powerful method for understanding the sources of variability in a functional data set.

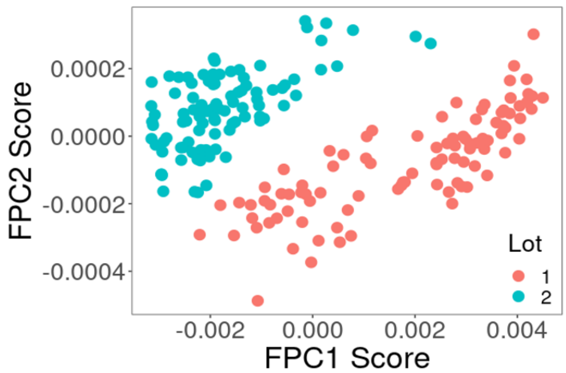

Figure 4: Functional principal component scores colored by production lot

D. Functional Regression

For a univariate or multivariate response variable, a standard statistical approach to investigating the effect of a factor on that response would be to use some form of regression model. There are several functional regression methods that can be used for the same purpose when the response variable is functional [3,4]. For our analysis, we use penalized flexible functional regression [5]. Our model is $C_i(V_{ij})=B(V_{ij})+B_0+B_1Lot_i+err_i(V_{ij})$ for device i and voltage value j. Here, $B(V_{ij})$ is the overall mean and B_0 is the intercept. Note that $Lot_i=1$ if device i comes from production lot 2 and is 0 if it comes from lot 1. With this variable encoding, $B(V_{ij})+B_0$ is the mean of the production lot 1 devices. We note that in this and other functional regression methods, the regression coefficients can be either scalar or functional. We explored both these options and found that using a functional coefficient did not add insight or improve the model in any meaningful way. For this reason, the only functional regression coefficient in our model is the intercept.

Table 1 contains the estimated intercept and

Table 1: Functional regression coefficient estimates with standard deviations and p-values

	Estimate	Standard Error	p-value
Intercept	-0.007	<0.01	<0.0001
Lot2	0.002	<0.01	<0.0001

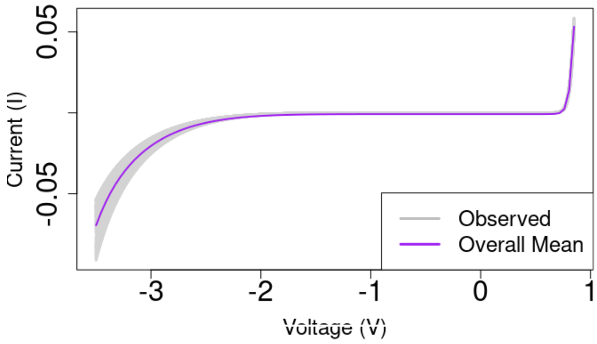

Figure 5: Full data set of I-V curves with overall mean from functional regression

production lot 2 coefficients along with their standard errors and p-values. These p-values are interpreted as in non-functional regression models. This model also has an adjusted R^2 value of 0.975, meaning that 97.5% of the variability in the data is explained by the factors we have included in our model. This indicates a very strong fit of the model to the data.

In Table 1, the intercept coefficient is interpreted as the amount of current the overall mean is changed by at each voltage value to arrive at the mean of the production lot 1 devices. Similarly, the Lot2 coefficient is the amount of current added to the production lot 1 mean to arrive at the production lot 2 mean. At the 0.01 level of significance, there is a statistically significant difference in the behavior of devices between production lots. Fig. 5 displays the full data set with the overall mean plotted on top of the I-V curves.

This regression analysis provides straightforward estimates for the mean device behavior in each lot – an important problem on its own [6]. More importantly for our analysis, we have quantified the effect of production lot on device behavior and found that it is statistically significant.

Conclusion

We have applied three FDA techniques to the analysis of a set of Zener diode devices from two production lots. These methods have uncovered insights about the effect of production lot on device behavior that were not evident from a visual inspection of the data. In future work we will use similar methods to analyze a data set of Zener diodes that have been subjected to an aging condition and measured over time in order to quantify the effects of aging on the devices.

In addition to the methods used here, there are a host of other FDA techniques that can be employed [3]. As many types of electrical device data are functional in nature, we advocate for the expanded application of FDA methods in this community. As demonstrated in this paper, such methods can provide principled quantitative analysis and useful insights for challenging data sets.

Acknowledgments

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & 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.

References

[1] Tucker, J. D., Wu, W., & Srivastava, A. (2013). Generative models for functional data using phase and amplitude separation. Computational Statistics & Data Analysis, 61, 50-66.

[2] Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. Chemometrics and intelligent laboratory systems, 2(1-3), 37-52.

[3] Ramsay, J. O., & Silverman, B. W., Functional data analysis, New York, Springer, 2005

[4] Reiss, P. T., Huang, L., & Mennes, M. (2010). Fast function-on-scalar regression with penalized basis expansions. The international journal of biostatistics, 6(1).

[5] Fabian Scheipl, Ana-Maria Staicu, and Sonja Greven. Functional additive mixed models. Journal of Computational and Graphical Statistics, 24(2):477–501, 2015K.

[6] N. Martin, T. Buchheit, and S. Reza, "Selection of a nominal device using functional data analysis," *Proc. - 2018 IEEE 5th Int. Conf. Data Sci. Adv. Anal. DSAA 2018*, pp. 381–386, 2019, doi: 10.1109/DSAA.2018.00049.