SAND2023-12394C

WTHPE-XXX (Poster ID)

Mean Estimation and Nominal Device Selection with the Pairwise Midpoint Method

Jason Adams, Thomas Buchheit, Abdullah Al Mamun Mazumder, Biazid Kabir Moghal, Md Fazle Rabbe, Ahsanul Islam, Shahed Reza

Sandia National Laboratories, MetroScientific, USA

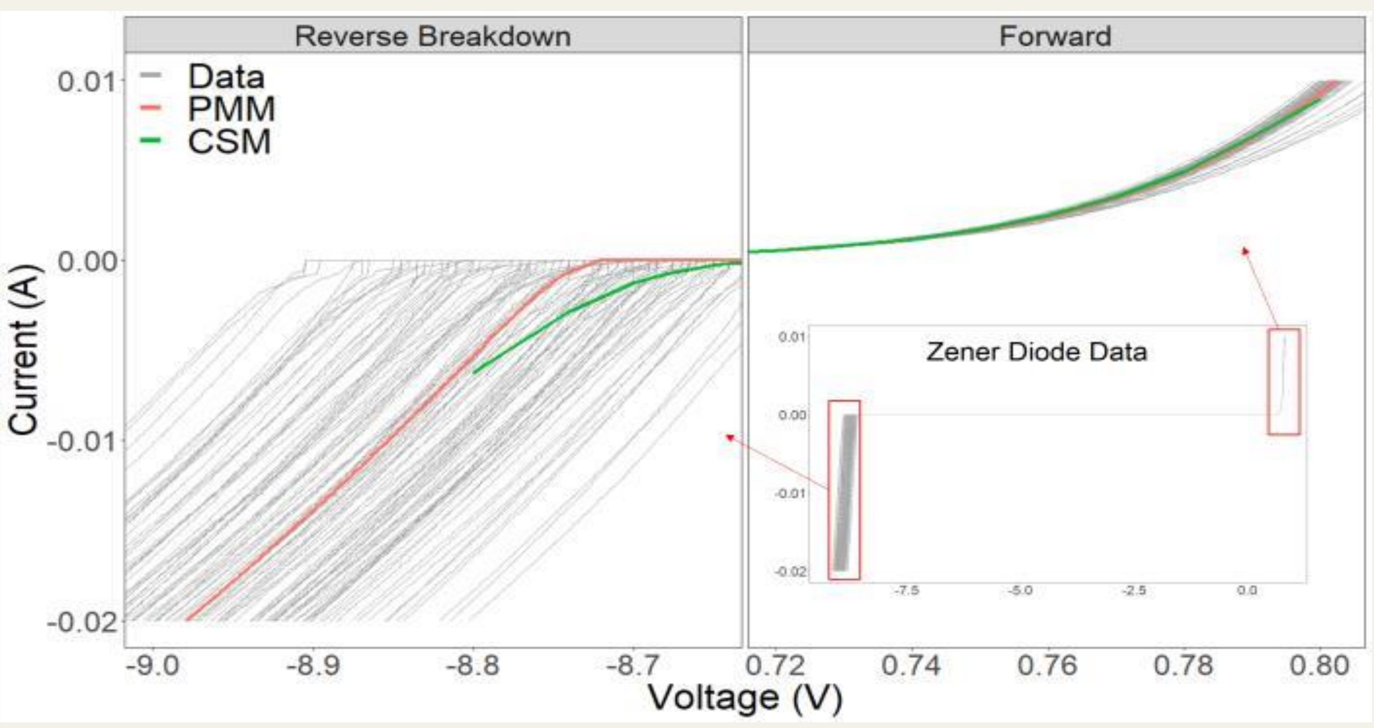
Motivation

- Accurate characterization of electrical device behavior is a key component of developing accurate electrical models and assessing reliability.
- The mean electrical behavior of a population of measured devices is valued for two reasons:
  - for use in circuit simulations
  - quantifies device variability, useful for reliability analysis
- Functional data analysis must consider amplitude (y-axis) and phase (x-axis) variability.
- Available methods:
  - **Cross-sectional Mean (CSM):** computed by averaging all observed function values at a given time point.
    - does not account for phase variability
    - poor representation of the mean behavior
  - **Karcher Mean (KM):** provides a much-improved representation of the mean compared to the CSM.
    - complex and time-consuming.

Pairwise Midpoint Method (PMM)

- Suppose there is a set of N functional observations at T time points. The data points can be denoted as $y_{ij} = f_i(t_{ij})$; $i = 1, \dots, N, j = 1, \dots, T_i$.
- Normalized in both y and t (if needed).
- Two observations are selected from N observations using the rule $y_{r1} < y_{s1}$; f_r = reference function, f_s = selected function.
- Let y_{rm} on f_r, and y_{sl} on f_s be determined when $|y_{rm} - y_{sl}|$ is minimized for all $l = 1, \dots, T_s$.
- Denoting the mean of f_r and f_s as $y^{rs} = f^{rs}(t^{rs})$, the m^{th} point of mean is computed as $((t_{rm} + t_{sl})/2, (y_{rm} + y_{sl})/2)$.
- Computing this for all $m = 1, \dots, T_r$ yields the mean function.
- This process is repeated for the remaining observations following pairwise selection, where each mean curve is weighted according to the number of original functions used to compute that mean.

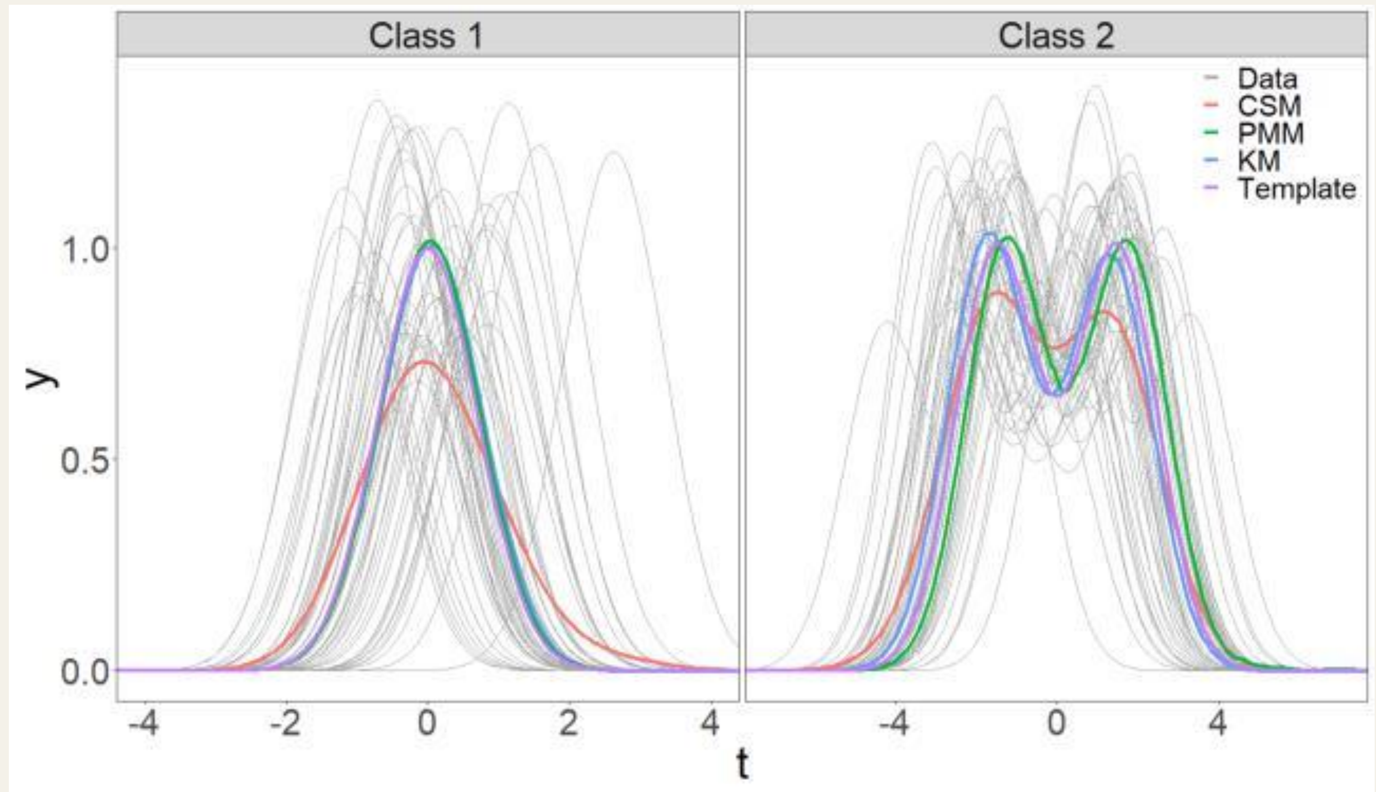
(a)



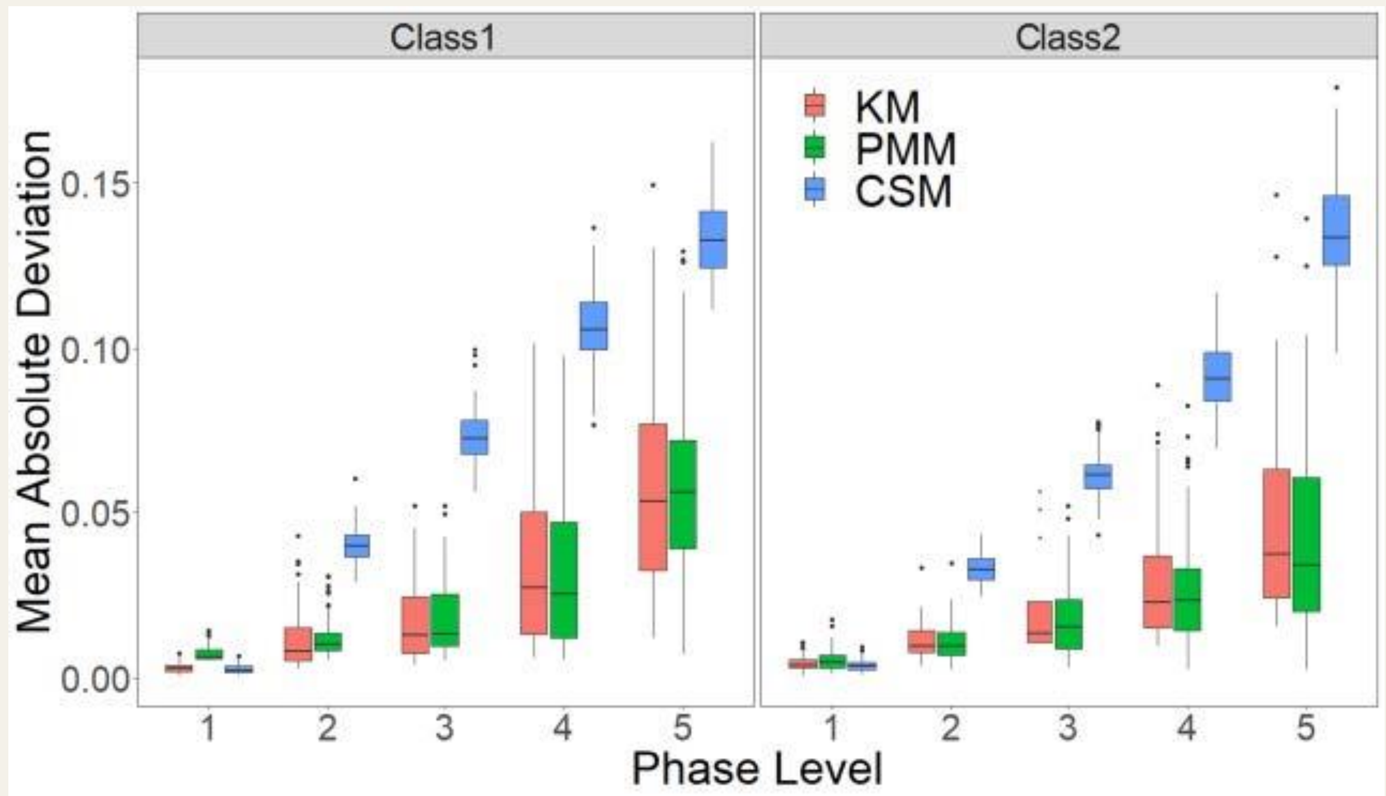
(a) Comparison of PMM and CSM in regions of interest for real data (144 MMSZ5239BT1-G Zener diodes).

Methods Assessment

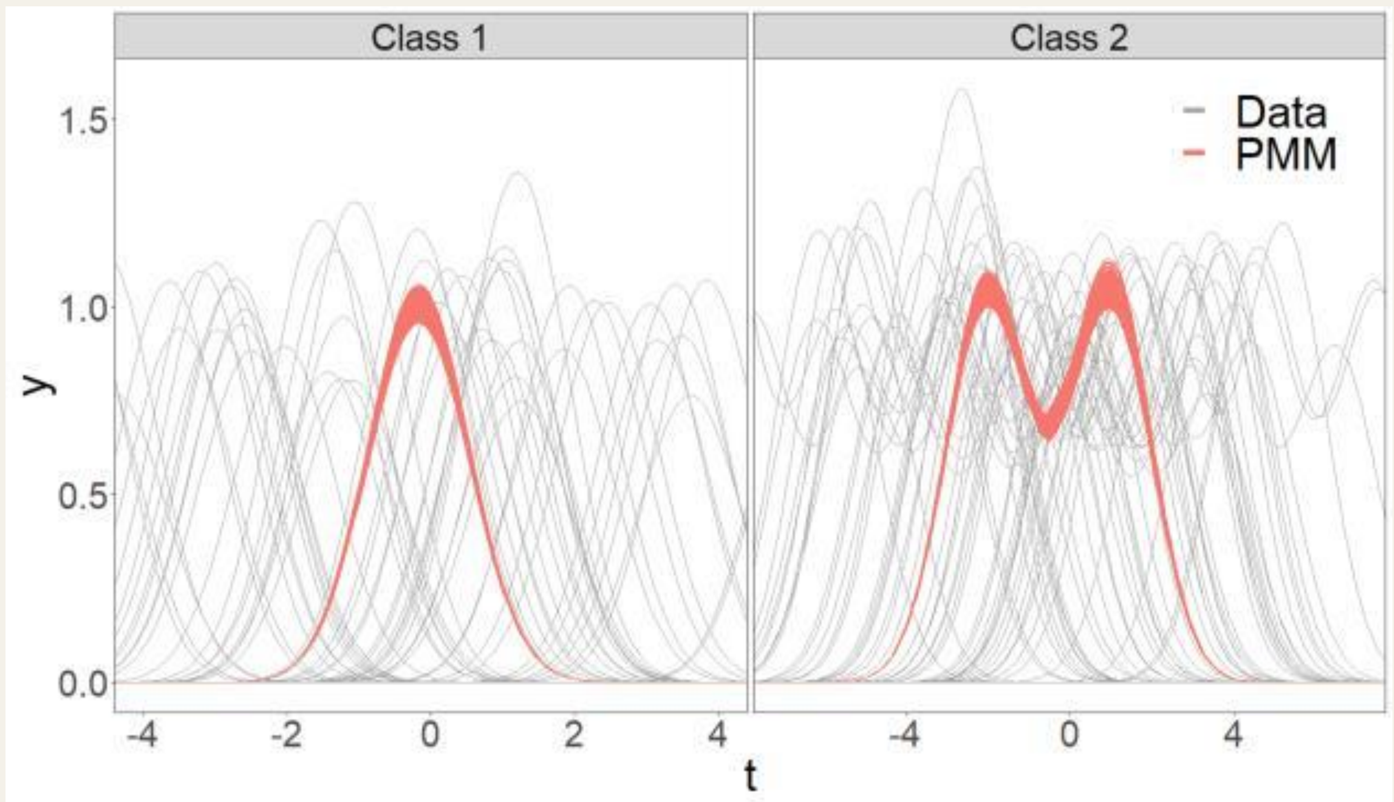
(a)



(b)



(c)



(a) Simulated data sets with means (CSM, PMM, KM) and template function
- Class 1: $a_i \exp[-1/2(t - p_i)^2]$,
- Class 2: $a_{i1} exp[-1/2(t + 1.5 - p_i)^2] + a_{i2} exp[-1/2(t + 1.5 - p_i)^2]$,
- Template: *setting* $p = 0$ *and* $a = 1$ *or* $(a1 = a2 = 1)$; *a (amplitude) and p (phase)*

(b) Mean absolute deviation (MAD) values for mean estimation methods

(c) PMM means for random variation of input sequences

Timing analysis

Method	Time (seconds)
CSM	< 0.001 (< 0.001)
PMM (Proposed)	**0.257 (0.011)**
KM	209.919 (0.580)

Conclusion

The numerical study demonstrated that PMM is a viable method for mean estimation as

- it performs well, even in the presence of phase variability (in contrast to CSM)
- it is much less computationally expensive than KM.

ACKNOWLEDGEMENT: 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.