



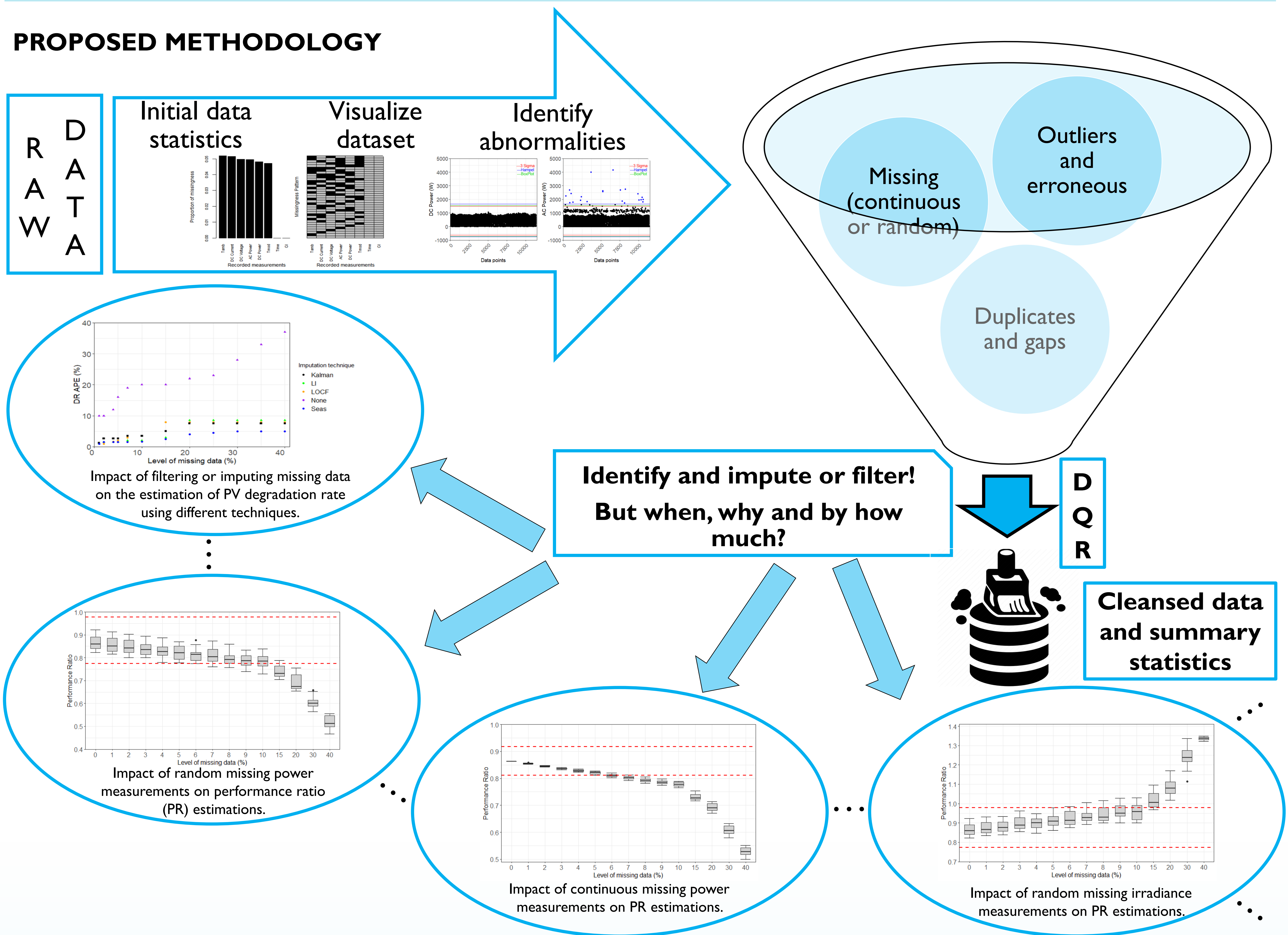
Guidelines for ensuring data quality for photovoltaic system performance assessment and monitoring

A. Livera (UCY), M. Theristis (SNL; mtheris@sandia.gov), E. Koumpli (SC), G. Makrides (UCY), J. S. Stein (SNL), G. E. Georghiou (UCY)

INTRODUCTION

- Invalid data (i.e. missing, erroneous and outlying values) caused by power outages or component failures are common problems exhibited in photovoltaic (PV) monitoring systems
- The way to handle such data can introduce bias in PV performance, reliability and degradation analyses
- A complete methodology of data quality routines (DQR) is being developed for data processing and quality checks
- Builds on IEC 61724 and other reports that provide justifiable and quantifiable criteria (mainly filtering, uncertainty, sampling) and also provides initial and summary data statistics
- Each step/decision is described in a quantitative manner based on detailed analyses and not arbitrary assumptions
- Potential of creating an open-source library for PV performance analysts and researchers

PROPOSED METHODOLOGY



GENERAL SUMMARY AND CONCLUSIONS

- DQRs are being developed to ensure data sanity and validity for PV reliability, monitoring and performance analyses
- Algorithms are used for identifying, imputing and/or filtering abnormal data (missing, erroneous, duplicate, etc.)
- It was verified that PV performance analysis is sensitive to the amount and location of abnormal data within a dataset
- Treating data using k-NN and seasonal decomposition proved to be effective
- A manuscript will be submitted soon and the complete procedure will be coded and published on GitHub