Testing the consistency of multimachine databases for physical studies of regression

A. Murari, M. Lungaroni and M. Gelfusa*Nucl. Fusion* **60** 015001;
https://doi.org/10.1088/1741-4326/ab4285

Abstract:
The investigation of various aspects of tokamak physics is performed with a combination of experiments carried out in different machines, to improve the statistical basis of the results and to cover a sufficient wide region of the operational space. Therefore, in the last decades, various multimachine databases have been built to address general and specific physical questions, particularly related to the extrapolation of present results to the next generation of devices. In this paper, a methodology of analysis is presented, to assess whether a multimachine data set is sufficiently coherent to really substantiate the conclusions, which are expected to be derived from it. A series of statistical and information theoretical tools have been refined to address the consistency of the data provided by the different devices. The developed techniques allow determination of whether it is reasonable to expect that the physics is the same in the various devices and/or that the entries do not present unacceptable bias. To exemplify the potential of the proposed approach, a systematic analysis of the ITPA database of the confinement time has been performed, using both dimensional and dimensionless quantities. The results obtained strongly suggest that better care should be taken in ensuring the coherence of data obtained from different experiments on different devices.-called transfection process.heat load environment are discussed.