A Big Wind Energy Data Analytic Paradigm - From PCA to Machine Learning[1]

Mo Jamshidi and Barney Tannahill[2]

ACE Laboratories

The University of Texas, San Antonio, USA

ABSTRACT

Large sets of data (numerical, textural and image) have been accumulating in all aspects of our lives for a long time. Advances in sensor technology, the Internet, social networks, wireless communication, cyber-physical systems (system of systems - SoS), and inexpensive memory have all contributed to an explosion of "Big Data". Big data comes in many ways in today's highly inter-connected world. Social networks, SoS and wireless systems are only some of the platforms creating big data. Wind energy has become a viable alternative to fossil fuels, especially in Europe. Recent efforts have developed a promising approach, called "Data Analytics", which uses statistical and computational intelligence (CI) tools such as principal component analysis (PCA), clustering, fuzzy logic, neuro-computing, evolutionary computation, Bayesian networks, etc. to reduce the size of "Big Data" to a manageable size. Then one can apply these tools to a) extract information, b) build a knowledge base using the derived data, and c) eventually develop a non-parametric model for the "Big Data". This paper attempts to develop a data analytic approach for wind energy forecasting. Database of wind energy are used here.

---

1. This work was supported, in part, by University of Texas System's Lutcher Brown Endowment, University of Texas at San Antonio.

2. With Southwest Research Institute (SwRI), San Antonio, TX, USA.