In this thesis, a data-driven testing procedure for wind turbine generators is developed. The procedure generates a data set for a hardware-in-the-loop testing setup at a test facility. The goal is to shorten validation process, prevent damage from highly dangerous grid tests, and conduct different tests. The proposed procedure deploys a multivariate statistical model of the power grid derived from the wind farm's standpoint. A practical data set of an operational wind farm is available which is logged during the years of 2013 and 2014 for model identification and validation, respectively. The first step of modeling is deriving the model from the standpoint of an arbitrary wind turbine generator utilizing dynamic principal component analysis. The model is the transformed data samples into a new projected space, i.e. latent space, as the combination of the principal components and the scores. Then, we propose an algorithm that optimally extends the derived model from the data set of a wind turbine generator to the wind farm. The optimized principal components result in the modeling error less than 5% and the selected scores cover the variance of the data with probability higher than 95% among all generators in the wind farm. As the next step, the joint probability distribution of the scores is estimated and then transformed into the sample space. Having the joint probability distribution, two approaches are proposed in order to generate testing data; off-line and on-line. In the off-line approach the data is generated for a time interval in advance while the on-line approach uses the current information to generate the next sample. For both approaches, an algorithm is implemented based on Gibbs Sampler to draw samples randomly given the joint probability distribution. Finally, using the off-line approach a testing data set is generated. Comparing the testing data with the validation data (year 2014) using energy function method proves the validity of the generated data.