本研究通过计算真实光伏出力数据集和生成数据集的概率密度函数（Probability Density Function，PDF），利用数据集的统计特性来评估生成数据的质量。首先，加载真实数据和生成数据，并将其展平成一维数组。概率密度函数的计算步骤如下：

1. 数据离散化：将数据分成若干个等宽的区间（bins），常用的区间数为100。设定的数据区间数为，每个区间的宽度为。

2. 直方图计算：对每个区间计算数据点的频数。记频数为，其中。

3. 频率归一化：将频数转换为概率密度，通过将频数除以总数据点数和区间宽度，即：

其中，为区间的中心点。

通过上述步骤，可获得每个区间中心点的概率密度值，进而绘制概率密度函数折线图，从而用于比较真实数据和生成数据的分布情况。这种方法可以直观地展示生成数据是否有效模拟了真实数据的统计特性。