作业五 海浪资料统计分析

**要求：**

1. 画出波面高度时间序列，利用上跨零点或者下跨零点的方法读取数据资料中的波高、周期等波浪要素，并画图展示。

2. 画出波高的概率密度分布，并计算给出有效波高和有效波周期

**数据使用：**

* data.txt
  + 一维海浪时间序列，共2048个数据，总时间长度为512s，时间间隔0.25s，数据单位为米（m）

**编写环境说明：**

3.10.10 | packaged by Anaconda, Inc. | (main, Mar 21 2023, 18:39:17) [MSC v.1916 64 bit (AMD64)]

numpy                          1.24.3

pandas                         1.5.3

matplotlib                     3.7.1

matplotlib-inline              0.1.6

scipy                          1.10.1

statsmodels                    0.13.5

seaborn                        0.12.2

**计算步骤**

1. 数据读取并生成时间序列
2. 对海浪波面观测资料进行平稳性检验，去掉趋势项
   * 平稳性检验
     + 其中，是时间序列经过差分后的系数估计值，是对进行标准化的估计标准误。若的值小于对应的临界值，则可以拒绝原假设，即可认为该时间序列是平稳的。
   * 线性回归法去除线性趋势
     + 线性回归法的基本思想是建立一个线性模型，将时间序列中的趋势部分与其他部分（随机部分）分离出来。具体地，假设原始序列为，可以建立一个线性回归模型：
     + 其中，是截距项，是斜率项，是随机扰动项。回归方程中即为时间序列的趋势部分，具有显著的统计意义。为了消除趋势，可以将减去和的值，得到线性趋势消除后的序列：
     + 消除了线性趋势后，序列的残差部分就是一个随机序列，可用于进一步的平稳性分析。
   * **结果：**
     + **ADF检验的p值为： 4.06\*10-29**
     + **去掉趋势项后的ADF检验的p值为： 4.99\*10-29**
3. 计算周期，波高
   * 有效波高是指一个海浪序列中，从所有波峰高度中选取最大的前1/3的数据点，计算其平均值
   * 显著波高是指一个海浪序列中，从所有波峰高度中选取最大的前1/10的数据点，计算其平均值
   * **结果**
     + **上跨零点法 1/10大波平均波高(显著波高)：1.447m 对应周期为：12.000s**
     + **上跨零点法 1/3 大波平均波高(有效波高)：1.079m 对应周期为：8.857s**
     + **下跨零点法 1/10大波平均波高(显著波高)：1.339m 对应周期为：11.906s**
     + **下跨零点法 1/3 大波平均波高(有效波高)：1.097m 对应周期为：9.152s**
4. 绘制概率密度
   * 核密度估计是一种非参数估计方法，用于估计随机变量的概率密度函数。核密度估计方法可以不需要事先对概率密度函数做出假设，可更精确地反映实际数据的分布情况。核密度估计的基本思想是对每个样本点，基于一定的核函数，构造一个形如：
   * 的非参数估计函数。其中，是核函数，是带宽参数。核函数通常是以为中心的对称函数，常用的核函数有高斯核、Epanechnikov核、直方图核等。带宽参数则决定了估计函数的平滑程度。核密度估计的核心在于估计非参数函数，它表示的是在处的概率密度函数估计值。核密度估计的优点在于可以从数据中得到概率密度分布，避免了需要进行假设检验等复杂的前提假设，同时，可以更加展现数据的分布情况，便于观察和分析。

**结果展示：**

图片包含 天线, 物体, 游戏机, 测量

描述已自动生成

Figure 1原始数据-时间序列

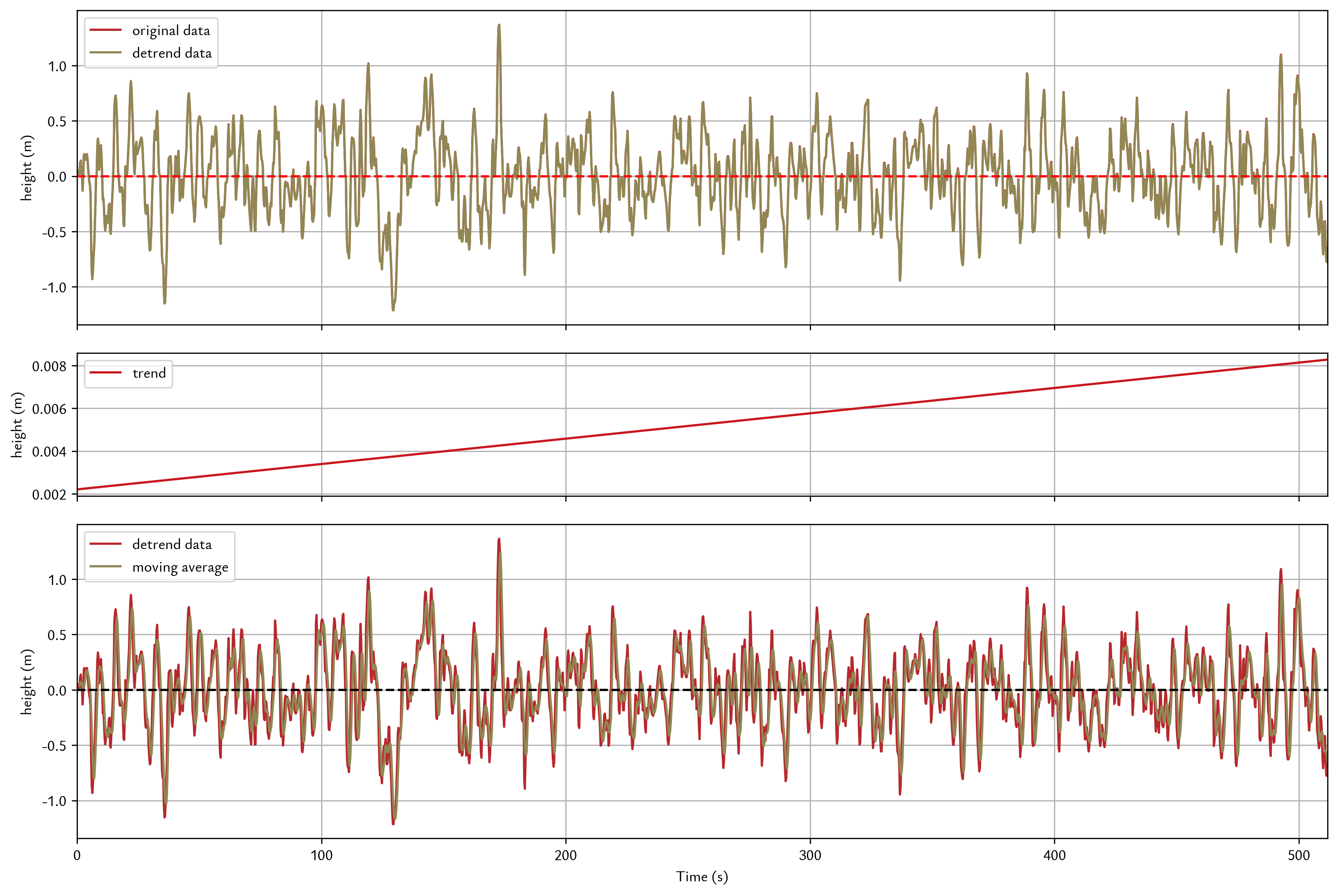


Figure 2 原始数据去趋势过程

图形用户界面, 散点图

描述已自动生成

Figure 3 上（下）跨零点位置

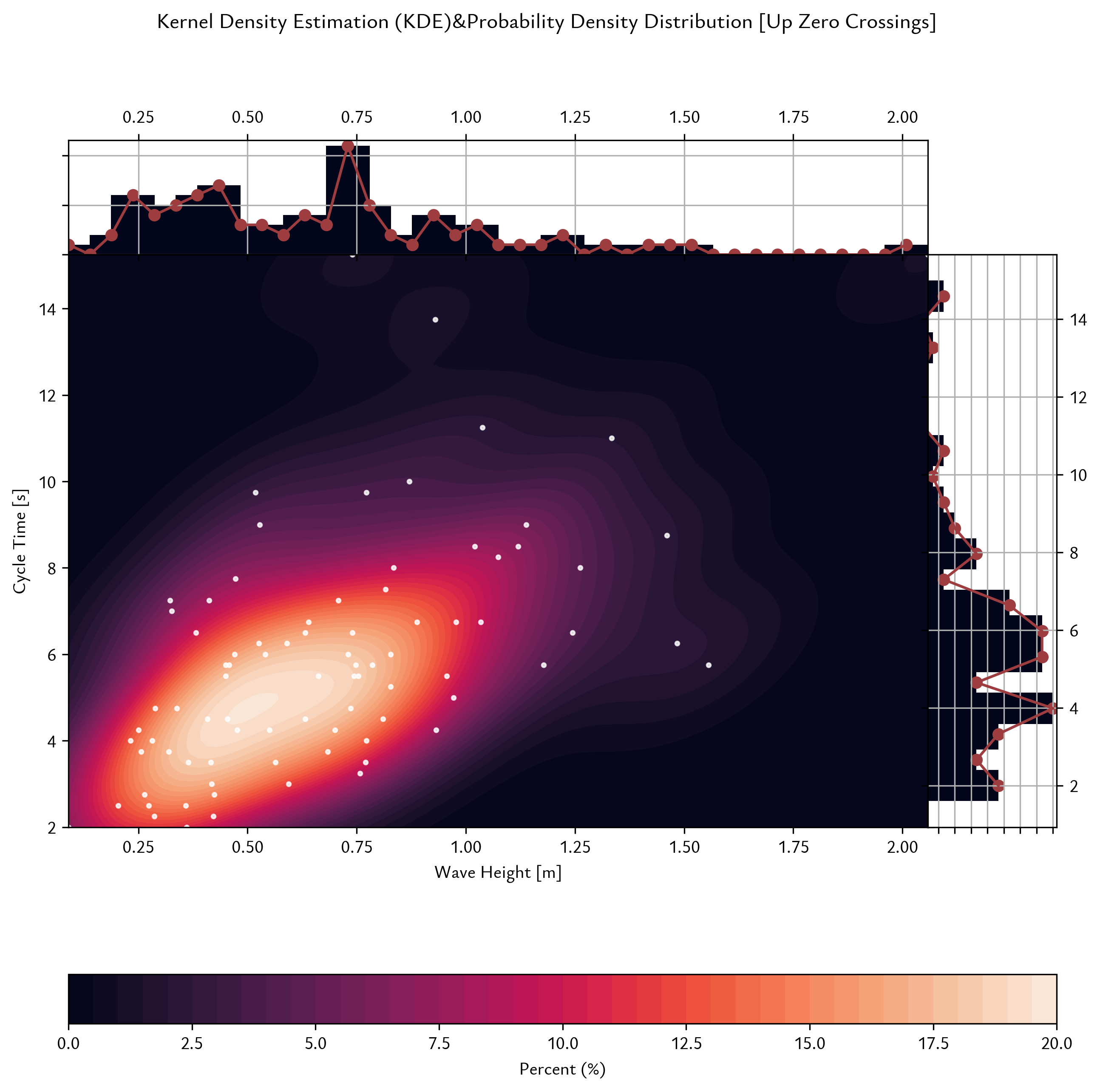


Figure 4 上跨零点法 概率密度分布 核密度分布图

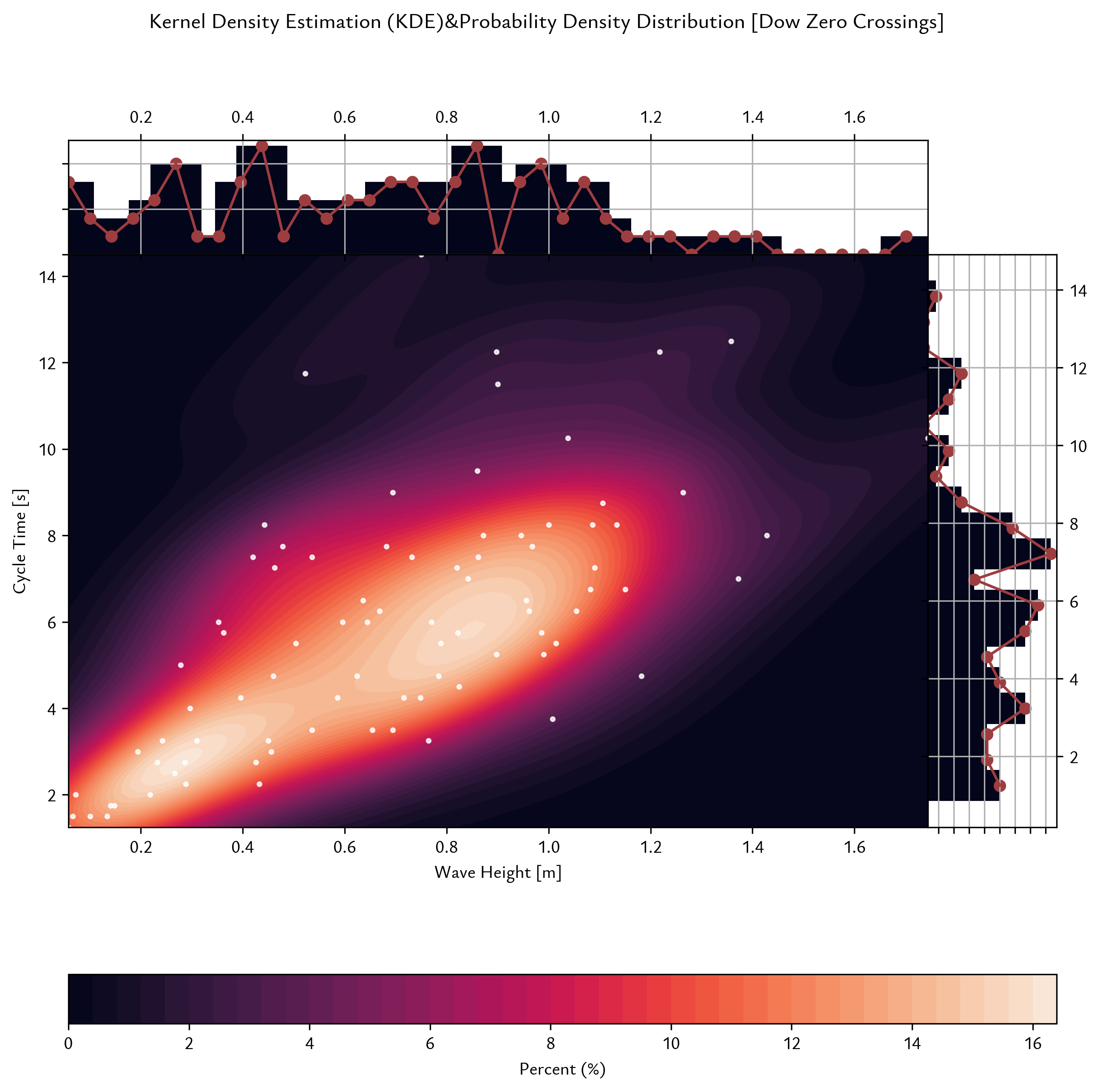


Figure 5 下跨零点法 概率密度分布 核密度分布图