

# 基于集聚点分析的东海岛北岸垃圾分布研究

**摘要：**海洋环境治理是一个值得深入研究的方向，随着社会经济的飞速发展，工业垃圾和生活垃圾也越来越多，由于垃圾从沿岸地区向海洋深处渗透，导致了海洋生态环境遭到了破坏。为了处理这些海洋垃圾，我们国家中的相关部门需要投入较大的成本进行处理，如何高效地对海上垃圾进行清理、回收，就成了一个亟待解决的问题。

本文以东海岛为例，实现了湛江湾近海垃圾聚集点的预估：根据沿岸水流对垃圾的运输作用、水流交汇点计算建立垃圾集聚点模型，通过模拟计算得出垃圾分布点；然后建立了基于水流与海岸距离的权重模型，为每个垃圾集聚点赋予合理的权重；再通过带权重的 K-Means 聚类分析预估垃圾聚集中心，此结果可以为相关部门提供垃圾据集中心位置，方便对海面垃圾的处理，为当地的海洋环境保护提供一个良好的参考。

**关键词** 集聚点模型 湛江湾 聚集中心预估

湛江湾具有非常重要的战略地位，海域曲折复杂且形状不规则，且南部有东海岛、硇洲岛等作为天然屏障，港内水深浪静，回淤少，四季不冻，是得天独厚的优良港湾。湛江湾作为湛江市区的战略资源，具有悠久的海洋渔业发展历史，经过十余年的扩大与发展，已成为湛江最为重要的海洋渔业养殖区，湛江湾中拥有大量的水产品，种类丰富，质量优越，常年作为广东省重要海鲜食品来源，湛江湾的战略意义不仅仅体现在经济发展上，湛江湾对湛江的气候调节也起着重要作用，相同的价值也体现在空气质量改善上——湛江湾丰富海洋生物的光合作用。除此之外，凭借海洋生态系统的自净功能，湛江湾每年还能处理大量污染物，这一点正是最宝贵同时也是最易被破坏的战略价值。湛江经济飞速发展的同时，也面临着潜在的环境问题<sup>[1]</sup>。如何在发展临湾重工的同时保护好湛江湾的生态环境价值，是值得所有人共同探讨的问题。

## 1 集聚点模型

由于垃圾在近海地区主要的输运动力是水力运输，可以认为，水流的交汇点也往往是垃圾的交汇点。于是可以根据沿岸地区每一点的水流对垃圾的运输与水流交汇计算从而得出垃圾分布点，从而建立如下的垃圾集聚点模型。

如图 1，假设有两条汇聚于点  $C$  的水流，设水流由  $S$  流向  $A$ ，在  $A$  处反射后流向  $C$  而形成反射水流  $AC$ ，同理假设存在反射水流  $BC$ 。

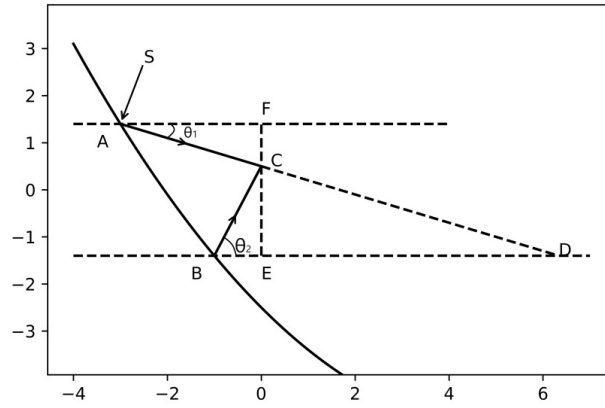


图 1 水流交汇示意图

假设点  $A$  为  $(x_1, y_1)$ , 点  $B$  为  $(x_2, y_2)$ , 水流的交点  $C$  为  $(x_3, y_3)$ 。由几何关系得:

$$\frac{CE}{BE} = \tan \theta_2 = \frac{y_1 - y_2}{x_3 - x_2}, \quad \frac{CF}{AF} = \tan \theta_1 = \frac{y_3 - y_1}{x_3 - x_1}, \quad \frac{CE}{DE} = \tan \theta_1 = \frac{y_3 - y_2}{x_3 - x_D},$$

又因为  $y_2 = \tan \theta_1 x_2 + b$ ,  $y_3 - y_2 = \tan \theta_2 (x_3 - x_2)$ , 有  $y_3 = \tan \theta_2 (x_3 - x_2) + y_2$ , 此时

$$\frac{y_3 - y_1}{x_3 - x_1} = \frac{\tan \theta_2 (x_3 - x_2) + y_2 - y_1}{x_3 - x_1} = \tan \theta_1,$$

即  $\tan \theta_2 (x_3 - x_2) + y_2 - y_1 = \tan \theta_1 (x_3 - x_1)$ , 于是得到如下垃圾集聚点  $C$  的计算式

$$\begin{cases} x_3 = \frac{y_1 - y_2 - \tan \theta_1 x_1 + \tan \theta_2 x_2}{\tan \theta_2 - \tan \theta_1} \\ y_3 = \tan \theta_2 (x_3 - x_2) + y_2 \end{cases}.$$

### 1.1 反射水流夹角

不难想见, 上述模型中我们需要计算反射水流与横轴的夹角  $\theta_1, \theta_2$ , 下面我们讨论该夹角的计算模型。由于海岸线总存在适当的截取与旋转使得海岸线在给定笛卡尔坐标下满足一一映射, 这样良好的性质能够方便计算并将作为讨论的前提。

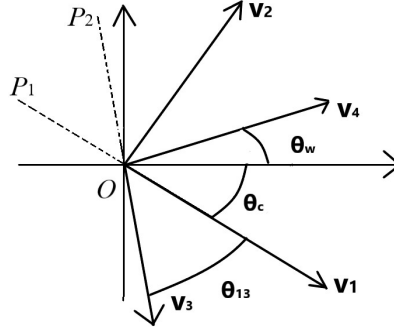


图 2 反射示意图

反射示意图如图 2，其中， $\mathbf{v}_1$  为海岸线切向量，所在直线即反射平面； $\mathbf{v}_2$  为垂直于  $\mathbf{v}_1$  的法向量。水流从  $P_2$  运动至点  $O$  反射的过程中， $\mathbf{v}_3$  是水流的方向向量， $\theta_c$  是  $\mathbf{v}_1$  和  $x$  轴之间的夹角， $\mathbf{v}_4$  为反射线的方向向量。记  $\theta_w$  为水流反射后和  $x$  轴正半轴的夹角。

不妨设海岸线切向量定义在一、四象限，由反射过程知， $\mathbf{v}_2$  和  $\mathbf{v}_3$  之间是入射角，记为  $\theta_\lambda$ 。 $\mathbf{v}_2$  与  $\mathbf{v}_4$  的夹角是反射角，记为  $\theta_\text{反}$ ，则， $|\theta_\lambda| = |\theta_\text{反}| > 0$ 。

于是我们不难得到在给出海岸线轮廓与入射水流时，反射水流夹角  $\theta_w$  的计算式。

$$\begin{cases} \theta_{13} = \arccos\left(\frac{\mathbf{v}_3 \cdot \mathbf{v}_1}{|\mathbf{v}_3| |\mathbf{v}_1|}\right) \\ \theta_w = |\theta_{13}| + \theta_c \end{cases}$$

同理我们能够计算出所有情况下的反射水流与横轴的夹角。

通过上述模型，当给出入射水流与海岸线轮廓时，我们能够计算出所有的垃圾集聚点，但是，可能由于任意两条不平行的水流必然存在一个交点，当交点在反射水流的反方向，则称为虚交点；且由于海岸线的存在，有些水流交点在陆地上，因此在运用模型商户模型前我们要对这些点进行虚交点的过滤和海岸约束判断。

## 2 水速下降与集聚点权重模型

不难想见，在假定水速恒定不变的情况下，垃圾集聚点可能出现在任意的位置，但是水流在运输的过程会存在能量的损失，垃圾集聚点也因此不能出现在任意远处。由于水流经反射后，能量逐渐损失<sup>[2]</sup>，流动越远越难维持自身运动状态，

其携带的垃圾也随着距离海岸的距离增大而减小，故各个水流和集聚点应当有各自的权重。于是不妨假设反射水流随着离开海岸的距离  $x_d$  增大水流的权重遵循

$W_w = -be^{ax_d} + c(a, b, c > 0)$ , 其图像如图 3 所示。

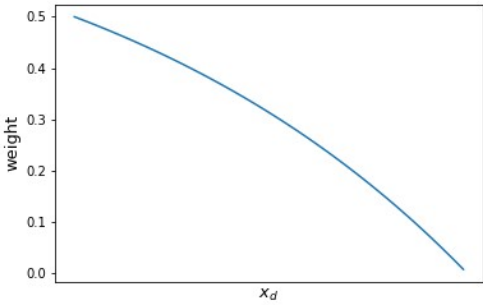


图 3 水流权重随海岸距离的变化曲线

则集聚点的权重为  $W = W_{w1} + W_{w2}$ , 即产生该集聚点的两条反射水流的权重之和。且总有构造满足  $x_d \rightarrow 0, W_w \rightarrow 0.5, W \leq 1$ .

### 3 集聚点中心预测

通过上述模型我们计算得到了垃圾集聚点和各集聚点的权重, 但这样得到的集聚点数量太大 (图 4), 且没有参考权重与密度, 参考意义不大。对于垃圾整治, 给出集聚的中心与参考位置便显得更为重要和可操作性。



图 4 湛江湾东海岛北部垃圾集聚点分布

现在考虑, 当给出所有可能的集聚点与各集聚点的权重后, 希望给出  $k$  个建议优先处理的重点区域。显然, 垃圾密度越高的地方越需要去处理, 越是处于局

部中心且权重较大的区域，越需要优先处理。而 K-Means 算法这一基于距离的聚类算法恰能为这个问题提供解决方案。<sup>[3]</sup>

K-Means 算法描述如下：

1. 指定聚类中心的数量  $k$  ,随机初始化  $k$  个聚类中心；
2. 计算每个点到各个中心的距离，距离最小的中心记为与该点同类；
3. 将每一类的点的加权平均更新为每一类的中心；
4. 重复 2-3 至每一类的中心不再发生变化。

显然，当我们选取欧几里得距离作为距离度量时，便能得到可供参考的  $k$  个集聚点中心。

### 3 东海岛北岸垃圾预测

在计算中，模型知道海岸线上每个点的水流方向，但这往往受限于观测有限，但是水流的主要驱动因素是风力，水流的宏观流向主要是由风向决定的，所以可以将风向作为水流的方向。



图 5 东海岛北岸垃圾集聚点及其中心（西北风）



图 6 东海岛北岸垃圾集聚点及其中心（东北风）

图 5、图 6 展示了不同风向情况下东海岛北岸垃圾集聚点分布的预测，其中，红星标记处是 K-Means 集聚点中心预测算法给出加权中心。从图中我们可以看到，当吹西北风为主的日子，垃圾在东侧集聚得更多；而当吹东北风为主时，垃圾往往随着水流远离东岸，漂流到了西侧。而这也说明了方法的合理性和有效性。但是从图中也可以看出，每个中心附近的集聚点密度是不同的，此时当地环保部分可以根据不同中心附近的密度进行排序来制定应对策略。

**总结：**随着国家海洋战略的推进，地处南海的湛江湾经济也不断发展，人口增加，但也意味着我们要更加重视湛江湾的环境保护。本文提出了一种基于水流输运的垃圾集聚点模型，并考虑水流流动的衰减特性，提出基于运动距离的水流与集聚点权重模型，为集聚点赋予合理的权重。权重参与 K-Means 集聚点中心预测的计算，得到集聚点群的加权中心。

经实验，该方法结果合理有效，得到的集聚点及其中心预测能够为当地保护部门提供有效的参考预测，方便海洋垃圾处理方案的制定。

## 参考文献

[1] 冯国艳. 湛江湾海岸带综合管理优化研究[D]. 广东海洋大学, 2019.

- [2] 林海卓, 杨哲, 辛博, 占海明. 基于阻尼振荡动态模拟环太平洋垃圾带的成因[J]. 科技传播, 2010(16):108-109.
- [3] 周艺. 基于中心聚集参数的改进 K-means 协同过滤推荐算法[D]. 河北经贸大学, 2020.