

塑料 Global Plastic Waste Model based on Linear Programming and Support Vector Regression

摘要

Plastic waste (PW) has been one of intractable environmental dilemma for modern humankind. In order to mitigate this tough problem, we combine production of plastics and management of plastic waste to establish a model, which can evaluate the impacts of plastic waste and find out an achievable minimal level of plastic production that can be reduced with a time line.

Firstly, we develop a plastic waste estimate model based on Product Life Cycle (PLC) theory to calculate the maximal levels of single-use or disposable plastic product waste within the environmental capacity in a certain region or country. Then we independently explore the four possible ends of PW and formulate their impacts on environment by several functions. We use linear programming that includes one particular objective function and four constraints. As a result, we can output the maximal amount of plastic products by analyzing the input digits of any region.

Secondly, to find out to what extent plastic waste can be minimized, we established an model called HSVR which combines happiness-index analysis and SVR method. By analyzing the characteristics of different regions and people' s living standards, we find out the relationship between the amount of plastics produced and people' s happiness index. Based on this, we can analyze how much plastic production we can reduce without significantly reducing happiness.

Thirdly, we follow the previous happiness index to assess how much human life is changed, and apply the Environmental Assessment of Solid Waste Systems and Technologies (EASEWASTE) model to evaluating how the the environment is affected by plastic reduction. Moreover, we take computable general equilibrium (CGE) model to quantify the effects of the target plastic production. Then we normalized the outputs and into non-dimensional indices, and calculate the weight for each indicator above by the analytic hierarchy process (AHP). The total impact for achieving the confined level is a linear combination of the three indices.

Furthermore, reduction of plastic production will impact countries unequally because of their different development level. We use dual programming to fine the dual model of the linear programming based model we established first. Then shadow prices of every factor formulated by constraints can be calculated, which indicate the significance of each factors. Then we divided them into environmental and governmental factors. Sensitivity analysis is be used to identify which factors that is most sensitive to a particular country, and the equitable reduction of plastic of the country depends on the categories that its sensitive factors belong to.

Finally, we analyze how to approach the minimum level following the time line by an achievable way. Many unexpected circumstance would definitely appear so we empirically choose most

significant possibilities that may delay or accelerate the achievement process. The ultimate result are specified into a memo which will be provided for ICM.

Keywords: Plastic waste; LCA; SVR; EASEWASTE; AHP

目录

1	Introduction	1
1.1	Background	1
1.2	Problem statement	1
2	Model Analysis	2
2.1	Analysis of task 1	2
2.2	Analysis of task 2	2
2.2.1	Rating-review based reputation model	2
2.2.2	Relations between time and reputation	2
2.2.3	Indication of potentially successful or failing product	2
2.2.4	relationship between specific star ratings and more reviews	3
2.2.5	relationship between specific quality descriptors of text-based reviews and rating levels	3
3	Assumptions and justifications	3
4	Parameter Table	4
5	Model Construction	5
5.1	Rating-review based reputation model	5
5.2	Reputation-time model based on RBF regression	6
5.3	Model to indicate potentially successful or failing product	7
5.4	Model to relationship between specific star ratings and more reviews	7
5.5	Model to relationship between specific quality descriptors of text-based reviews and rating levels	9
5.6	Theories related to our model	11
5.6.1	PLC theory	11
5.6.2	Linear programming	11
5.7	Qualify the impact on the environment	12
5.7.1	Producing process	12

5.7.2	Managing process	12
5.8	Constraints given by environmental carrying capacity	14
5.9	Model paradigm	14
5.10	Analysis of China using our model	15
6	Plastic waste reduced	16
6.1	SVR algorithm	17
6.2	Detailed analysis	18
6.2.1	The quantitative standard of happiness	18
6.2.2	Plastic Emphasized	19
6.3	Model parameter selection	19
6.3.1	Model performance evaluation	19
6.4	Specific model parameters	20
6.5	Predicting minimum plastic usage	20
7	The minimal target of PW and its impacts	22
7.1	Basic Theories	22
7.1.1	EASEWASTE	22
7.1.2	CGE	22
7.1.3	AHP	23
7.2	Global Target	23
7.3	Impact Path Analysis	23
7.3.1	How human life will be alert	24
7.3.2	How environment will be affected	24
7.3.3	How to influence plastic industry	24
7.3.4	Total impact	25
8	Discussion about equity	25
8.1	Shadow price	25
8.2	Determining shadow prices using dual theory of linear programming	26
8.3	Shadow price of China's plastic constraint equation: an example	27

9	Strength and weakness	27
9.1	Strengths	28
9.2	Weaknesses	28
10	Conclusion	28

Memorandum

To: ICM officials

From: Team 2013573

Subject: Global Plastic Waste Issue

Date: 2020 年 3 月 9 日

It is a great honor for us to be hired to address this increasingly severe situation crisis. Current global plastic waste has exceeded the environmental tolerance level, therefore an achievable level of reducing plastic waste is badly needed. Considering the mutual influence between plastic industry and human community, we have managed to find out an appropriate global target minimum of single-use plastic waste and the timeline to meet it.

We are pleased to tell you that we make it to estimate an appropriate target for a specific country. In Section4, the impact of plastic usage on human living standards is explored by SVR algorithm. And based on the thought of not affecting much of human living standards, training results of different SVR kernel functions are utilized to predict minimum plastic usage in a particular area.

As for the global target minimum achievable levels of single-use or disposable plastic product waste that we both concern most, we divided the world into 10 typical countries according to their plastic consumption amount, and for each category of countries the target for minimum plastic use is calculated. Then add up those predicted outcomes and a global target minimum achievable target is obtained. As we demonstrate in our paper, the total minimal plastic consumption obtained is about 6.6 hundred million tons.

Every target needs a schedule to realize, no mention to this grand plane to make the global plastic production and usage reach the defined level. Hence we need design a feasible timeline to achieve this level. However, according to our sensitivity analysis results, countries have various sensitivity to every single factor. It seems unrealistic to set the same timeline for all countries because they are in different industry structures. So the better way is to consider several timelines separately for disparate countries.

The majority of developed country has developed mature services. Although they use a lot of plastics, their plastic waste have been managed well due to their advanced recovery technologies. For them, reducing plastic means higher quality and efficiency of management. This is not a difficult task, but it requires a lot of perseverance and determination. However, manufacturing takes a prior part of domestic economy in most developing country and reducing the production of plastic is likely to threaten economy growth. For developing countries, reducing the use of plastic is a pain, because it means slowing development. In order not to deprive of the right of development of poor

population, we suggest a loose timetable for medium-income and low-income countries to optimize their industrial structure and reduce plastic waste.

Given all above we discussed, we decide to propose the timeline as follow:

By 2035, developed countries will first achieve reductions to established targets, and will begin to study how to further reduce the use of plastic. By 2050, all countries around the world will reach a lower level of plastic use.

The main consideration of this timeline is that the amount of plastic used in developed countries is much larger than that in developing countries, but they pay less to reduce it than in developing countries, as is presented before. But this issue cannot be generalized. For example, for the case we studied in China, the biggest problem is actually improper management. However, environmental and industrial factors cannot be ignored. The best way is to use our model to study all regions of the world as closely as possible to get an exact answer. This may be a laborious task, but it is definitely worth it.

There is a long way to go to achieve this goal between now and the future. So we list some circumstances that may affect the achievement of the target:

As for technology, if relevant research about plastic waste managing make remarkable progress and the findings are quickly implemented by environmental sector, the process of plastic reduction will be accelerated.

Besides, taking economic growth into account, it is likely that the government loosen regulations of plastic management in a recession to guarantee the economic growth. Take Bolsonaro administration for example, local government usually attach less importance to environmental issues.

Wish the given advice above would contribute to your work!

1 Introduction

1.1 Background

The plastic industry can date back to 1950s. Over the past 60 years, the production of plastic has grown rapidly, which has surpassed most other artificial materials[1]. While people then did not realize the potential negative influence of plastic usage to ecosystem, plastic waste gradually accumulated in the environment because of its non-biodegradable trait. Some scholars believe that plastic bags and Styrofoam containers can take up to 1,000 years to decompose[2]. 6,300Mt plastic waste has been generated in 2015, while only 9% of which had been recycled and reused. 12% of them were incinerated and the percentage of plastics that discarded in landfills or natural environment was up to 79%[1]. Plastics in nature especially in the ocean will cause a series of ecological problems. Apart from the chemical effects on organisms, plastic ingestion and entanglement are also threatening the diversity of species[3]. It is easy to imagine that if no measure is taken, humans will face severe degradation and pollution caused by enormous plastic waste.

Besides, the management of the plastic waste is one of the hardest issues on integrated municipal solid waste. There are many researches that focus on the plastic recovery routes based on the life cycle assessment approach (LCA) [4], and a few methods to assess the impact of solid waste system and technologies, which aim to develop more effective ways to migrate the plastic waste problem[5].

However, it seems that there are few researchers that explored the valuation model of estimating the maximal use of plastics nor utility policy to reduce the usage of plastics remarkably. That indicates there is still room for further explanation in this area.

1.2 Problem statement

Living in the age that are characterized by a series of troublesome environmental problems, such as global warming, desertification, pollution of water and poor air quality. Our team is hired to estimate the maximal permitted amount of plastic products, migrate the plastic production and reduce plastic waste. In another word, we need to answer the following questions:

1. What is the maximal level of single-use or disposable plastic products without further pollution?
2. What extent can plastic waste be reduced to?
3. What is the minimal achievable level of global single-use or disposable plastic products and what are the impacts on certain facets of reaching this level?
4. How to solve the equity problem about plastic waste that arises from global crisis?
5. How to build a timeline to approach our target level we set before and describe our exploration to ICM?

2 Model Analysis

2.1 Analysis of task 1

2.2 Analysis of task 2

2.2.1 Rating-review based reputation model

我们可以从产品的星级和评论推断出产品的声誉，而一个产品的声誉越好往往意味着产品的潜在销售量越高，在这里，我们不能单纯地用星级或评价衡量一个产品的声誉，因为生活中往往存在着“五星好评”和“一星黑粉”的情况。因此我们综合运用星级和评论构造出声誉的模型，通过对比产品的声誉，阳光公司可以针对性提升声誉高的产品的备货量和宣传力度。

同样，声誉的计算需要先对某一产品的评论进行 LDA 处理，转化为离散的等级值，为了考虑评论的可信度，我们还考虑了 vine, verified_purchase, helpful_votes, tot_votes 几个字段，这些字段可以对评论的准确度进行衡量，因为购买过产品的评论才更可信，helpful_votes 高的评论才更具有参考性。

通过建模，并对 xxxx 年 xx 月的 pacifer 样本 (xxx,xxx,xxx,xxx) 计算，我们发现，在该月内，样本 xxxx 的声誉更高，而 xxxx 的评论数也确实是最多的，说明我们的模型确实是可靠的。

2.2.2 Relations between time and reputation

产品的信誉会随着时间的变化而波动，而其中潜在的规律可能会对阳光公司有很大的启发。我们可以通过 RBF 神经网络对 (信誉-时间) 作函数拟合，在得到拟合曲线后，就可以直观地观测到信誉的变化趋势。以奶嘴为例，

- 首先，我们挑选奶嘴中评论数最多的样本-'philips avent bpa free soothie pacifier, 0-3 months, 2 pack'(样本数为 833)
- 然后利用信誉转换模型得到所有样本的信誉
- 接着以一天为单位时间，利用 RBF 神经网络，对信誉与时间作拟合
- 最后，根据拟合曲线，我们发现 philips 奶嘴的信誉总的变化趋势是 xxxx，在 xxxx 年 xx 月 xx 日最低，xxxx 年 xx 月 xx 日达到最低

对其他两类产品也可以作类似的处理，这些结果提示阳光公司 xxxxxx

2.2.3 Indication of potentially successful or failing product

正如上文中提到的信誉会随时间波动，那么我们拟合的曲线自然会有拐点。因此我们定义如果曲线将迎来极大值，那么这个产品将成为潜在的成功产品；如果曲线将迎来极小值，那么这个产品将成为潜在的失败产品。

为了方便处理，我们建模时只考虑曲线的最大和最小值，可以通过对曲线最大、最小值附近的星级、评论进行聚类，找到词频最高的主题词与星级。当这些主题词和星级出现的时候就意味着很

可能产品的声誉将迎来拐点。

在我们的实验中发现当 xxx 和 xx 出现的时候, 产品的声誉将下降, 那么这个产品会成为潜在的失败产品; 当 xx 和 xx 出现的时候, 产品的声誉将上升, 这个产品会成为潜在的成功产品。

2.2.4 relationship between specific star ratings and more reviews

特定的星级是否会引起更多的评论, 对此我们可以判断这是一个统计学问题。我们需要统计前一时间段内的平均星级, 以及下一个时间段内的评论总数, 来寻找它们的关系。以 Pacifier 为例

- 首先, 我们挑选 Pacifier 中评论数最多的样本-‘philips avent bpa free soothie pacifier, 0-3 months, 2 pack, packaging may vary’(样本数为 833)
- 然后统计该产品的不同时间段内的评论数, 以及时间段之前的平均星级。
- 因为自变量只有星级, 我们将其离散化为高星级, 中等星级, 低星级。接着我们使用单因素方差分析, 寻找星级与评论之间的关系。
- 最后, 根据统计结果, 我们发现在当前时间内的平均星级与下一个时间段内的评论总数间关系不强, 但是经过对评论语义分析, 我们将评论分为分为好评, 中评, 差评。我们发现在面对一段时间内产品都是中等星级时, 用户更愿意给出差评和中评; 在面对一段时间内产品都是高星级时, 用户更愿意给出好评。

2.2.5 relationship between specific quality descriptors of text-based reviews and rating levels

题目要求我们寻找出现在评论中的特定的词, 如”disappointed”, ”enthusiasticer”, 是否会和评论的等级有关系。因为是寻找数据间的关系, 我们可以判定这还是一个统计学问题。以吹风机为例

- 首先, 我们对吹风机中的评论进行词频统计, 选出频率较高的, 且是 specific quality descriptors 的词语。
- 然后, 我们挑选吹风机中含关键词的评论, 我们基于是否含关键词, 对评论进行 0,1 编码。
- 接着因为自变量只有是否含关键词, 所以我们继续使用单因素方差分析, 寻找关键词与星级之间的关系。
- 最后, 根据统计结果, 我们发现含关键词 “disappointed” 的评论和星级 1 密切相关, 含关键词 “love” 或 “great” 或 “good” 的评论和星级 5 密切相关。

3 Assumptions and justifications

To simplify the problem, we strive for the following assumptions:

- The mass of annual plastic production equals to that of the annual plastic waste. On the one hand, every pound of plastic produced will turn into waste. On the other hand, the lifetime of most sorts of plastic is under 20 so that given situation will not change acutely[1].

- Throughout the life cycle of plastic, the impact of plastic on environment in use phase is negligible. Main environment issues that are related to plastic merely when it is produced or managed[11].
- Aiming to make plastic waste safely be mitigated, any waste should not be landfilled. While it is extremely hard for plastic to biodegrade in solid, landfills are in great shortage.[11].
- All the plastic waste produced will be managed. In the long run, the NET plastic waste discarded in natural environment will approach zero by governments and NGOs' efforts.
- The proper quantity of pollutant emission of plastic waste should depend on the contribution of plastic industry to the gross domestic products. The pollutant emission of plastic waste is just an ordinary part of the total emission, so the responsibility of environmental protection it takes should also be the proportional part, which is roughly estimated by the percentage of GDP that plastic industry contributes.
- All the plastic waste is mechanically sorted from plastic waste prior to incineration. In fact, waste incineration plants are generally very formal and will complete the sorting work. Otherwise, they will be banned soon.
- The material recycling facilities(MRFs) that sort and upgrade the received waste stream is the only way to recycled sorted PW. There is no other mature processing method, and this method is used almost everywhere.

4 Parameter Table

Important symbols that in this article are listed in Table 1.

表 1: Notations

Symbols	Definition
s_i	Annual certain plastic production.
N_i	The number of CO_2 generated by the combustion of 1 molecule specific plastic unit.
EA	Emission to air.
EW	Emission to water.
EC	Emission of CO_2 .
TEA	Total emission to air.
TEW	Total emission to water.
TEC	Total emission of CO_2 .
CEA	Compensatory emission to air of recovery process.
CEW	Compensatory emission to water of recovery process.
SM	Plastic waste into marine.
M	The quantity of annual management of marine plastic.
V_{air}	Total available air of a specific country/region.
V_{water}	Total available water of a specific country/region.
σ	Proportion of incinerated plastic waste.
μ	Compensate.
θ	The percentage of GDP plastic industry contributes, quantified by coefficient.
ι	Tossil fuel consumption percentage. Generally, ι equals 4%.
k	Proportional coefficient of carbon emission per unit of energy consumption

5 Model Construction

5.1 Rating-review based reputation model

通过对评级和评论建模可以得出产品的信誉，而更高的信誉意味着更高的潜在销量，可以帮助阳光公司作出更好的决策。建模过程如下：

假定产品的星级为 R_α ，评论量化之后为 R_β ，评论的 helpful votes 为 R_h ，则 unhelpful votes(R_u) 为 total votes 减去 helpful votes，净投票数为 $(R_h - R_u)$ ，我们用净投票数加 1 作为因子再乘上量化后的评论，表征评价的可信度。

我们采用对数形式使得离散的各项参数值更平滑，信誉模型初步建立为：

$$F_{repu} = \log_2(R_\alpha + 1) + \beta \times \log_2(R_\beta \times (R_h - R_u + 2)) \quad (1)$$

更进一步，超参数 β 一方面要能体现评级和评论的权重关系，另一方面还要能表征评论的可信

度, 我们认为 vine 字段和 verified_purchas 字段为 'Y' 的评论是体验过产品的评论, 更具有参考性。因此 β 的表达式如下:

$$\beta = \begin{cases} xx & \text{if } vine \text{ or } verified_prurchase = 'Y' \\ 1 - xx, & \text{else} \end{cases} \quad (2)$$

其中 xx 为上文中提到的 xxx 关系系数。为了验证模型的准确性, 我们对 xxxx 年 xx 月 xx 号的所有 pacifier 产品做了声誉值计算, 结果如图:

然后我们又统计了该月 pacifier 各类产品的 vine 或 verified purchase 评论数近似作为该月的销量:

图中说明, pacifier 中的 xxxx 产品声誉值更高, 其在该月的潜在销量也越高。

5.2 Reputation-time model based on RBF regression

探究产品信誉随时间的变化规律有助于阳光公司调整策略, 增加收益。此处我们用 RBF 神经网络作信誉-时间的回归拟合, 因为 RBF 具有很强的非线性拟合能力, 可映射任意复杂的非线性关系, 而且学习规则简单, 便于计算机实现。我们的 RBF-based reputation-time model 建立过程如下:

1. 我们先对数据进行筛选, 选择每类产品最具代表性的样本:

- Pacifier: 'philips avent bpa free soothie pacifier, 0-3 months, 2 pack, packaging may vary', 有 833 个样本。
- Hair Dryer: 'remington ac2015 t|studio salon collection pearl ceramic hair dryer, deep purple', 有 587 个样本。
- Microwave: 'danby 0.7 cu.ft. countertop microwave', 有 394 个样本。

2. 将筛选出的样本代入信誉模型, 得到样本信誉随时间变化的离散序列。

3. 搭建 RBF 网络:

- (a) 将样本的前一半作为训练集、后一半作为测试集, 并作归一化处理
- (b) 设定网络的神经元数逐步增加, 最多就是训练样本个数, 扩展速度 spread 设定为 0.1, 开始训练
- (c) 测试 RBF 网络, 得到拟合结果图。

我们最后得到的拟合结果图如下所示:

从图中我们可以得出以下结论:

- Pacifier 总的趋势是 xxxx, Hair Dryer 总的趋势是 xxx, Microwave 总的趋势是 xxx, 这可能是因为 xxxxx

- 对于 Pacifier 信誉最低出现在 xxxx 年 xx 月 xx 号, Hair Dryer 信誉最低出现在 xxxx 年 xx 月 xx 号, Microwave 信誉最低出现在 xxxx 年 xx 月 xx 号。具体的原因分析我们将在下一问中进行阐述。
- 对于 Pacifier 信誉最高出现在 xxxx 年 xx 月 xx 号, Hair Dryer 信誉最高出现在 xxxx 年 xx 月 xx 号, Microwave 信誉最高出现在 xxxx 年 xx 月 xx 号。同样, 我们将在下一问讨论出现该现象的原因。

5.3 Model to indicate potentially successful or failing product

在我们的建模中, 拟合曲线最大值附近常出现的主题词和星级是出现潜在成功产品的标志, 而拟合曲线最小值附近常出现的主题词和星级是出现潜在失败产品的标志。

以 Pacifier 为例, 根据图??, 拟合曲线最大值出现的时间点为 xxxx 年 xx 月 xx 日 (标记为 t_1), 拟合曲线最小值出现的时间点为 xxxx 年 xx 月 xx 日 (标记为 t_2), 分别抽取 t_1 , t_2 附近 7 个时间点的星级和评论, 我们进行 LDA 文本分析并聚类, 结果如下:

我们发现, 最大值附近, xxx 和 xx 容易出现, 这很可能是因为 xxxx; 最小值附近, xx 和 xxx 容易出现, 而这可能是因为 xxx。

5.4 Model to relationship between specific star ratings and more reviews

在我们的建模中, 因为有且只有特定的星级会引起用户更多的评论, 显然只有星级一个自变量, 所以我们采用单变量方差分析, 我们的 Model to relationship between specific star ratings and more reviews based on One-Way ANOVA 建立如下

1. 我们先对数据进行筛选, 选择每类产品最具代表性的样本:

- Pacifier: 'philips avent bpa free soothie pacifier, 0-3 months, 2 pack, packaging may vary', 有 833 个样本。
- Hair Dryer: 'remington ac2015 t|studio salon collection pearl ceramic hair dryer, deep purple', 有 587 个样本。
- Microwave: 'danby 0.7 cu.ft. countertop microwave', 有 394 个样本。

2. 进一步筛选样本, 去除评论和星级不符的评论。

3. 然后我们将时间分段, 以 d 天 (商业一般为 7 天) 为一个周期, 假设在下一个周期的评论与上一个周期的平均星级有关。接着, 统计第 i 种产品的在第 j 个时间段的平均星级 \bar{s}_j , 以及下一个时间段的评论数 C_{j+1} 。我们需要挖掘 \bar{s}_j 与 C_{j+1} 的关系: 平均星级可以为 $[1, 5]$ 之间任意值, 为了方便寻求关系, 我们将其离散化, 可以通过将 $[1, 5]$ 均分为 4 个小区间, 每个小区间对应为一类, 再将 \bar{s}_j 根据其大小分到不同的类别 A_p 中。所以, 问题被转化成特定的星级区间 A_p , 是否对应更多的下一个时间段的评论数 C_{j+1} 。由于只考虑星级对评论数的影响, 我们采用单因素方差分析。

单因素方差分析要求各变量服从正态分布，且满足方差齐：

从图中我们发现满足方差齐，可以运用单因素方差分析：

表 2: Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
2.054	2	812	0.129

从图中我们可以发现

表 3: ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	12.31	2	6.15	16.599	0.001
Within Groups	13402.85	812	16.51		

- 在当前时间内的平均星级与下一个时间段内的评论总数间关系不强

后来，我们又对评论进行了语义分析：

1. 我们将评论根据语义分析的结果，分为了正面评价，中等评价，负面评价。
2. 然后我们统计时间段内正面平均，中等评价，负面评价的占比与 A_p 的关系

我们最后得到的统计结果图如下所示：

表 4: Descriptive of the percentage of negative comment

rating	N	Mean	Std.Deviation	Std.Error	F	Sig.
2	2	56.25	8.83	6.25	16.599	0.001
3	18	36.64	42.52	10.05		
4	175	8.68	19.47	1.47		

表 5: Descriptive of the percentage of medium comment

rating	N	Mean	Std.Deviation	Std.Error	F	Sig.
2	2	37.50	17.68	12.50	0.40	0.671
3	18	20.40	27.16	6.40		
4	175	22.32	25.52	1.93		

从图中我们可以得出以下结论：

- 在面对一段时间内产品都是低等星级时，用户更愿意给出负面评价。
- 在面对一段时间内产品都是高星级时，用户更愿意给出正面评价。

表 6: Descriptive of the percentage of medium comment

rating	N	Mean	Std.Deviation	Std.Error	F	Sig.
2	2	6.25	8.83	6.25	9.828	0.001
3	18	42.95	40.88	9.64		
4	175	68.99	29.36	2.21		

5.5 Model to relationship between specific quality descriptors of text-based reviews and rating levels

要判断是否在评论中出现关键词和星级是否有关，显然只有是否出现关键词一个自变量，所以我们仍然采用单变量方差分析，我们的 Model to relationship between specific quality descriptors of text-based reviews and rating levels based on One-Way ANOVA 建立如下

1. 我们先对数据进行筛选,统计出现频率较高的 specific quality descriptors,如”disappointed”,”love”,”great”.
2. 进一步筛选样本，因为有人会在 ‘disappointed’ 前加 ‘not’，使意思相反，所以去除这些样本
3. 基于评论是否含关键词进行热编码分类，并基于此进行单因素方差分析。

我们最后得到的统计结果图如下所示：

表 7: ”disappointed”

	N	Mean	Std.Deviation	Std.Error	F	Sig.
0	11250	4.16	1.27	0.012	745.936	0.000
1	220	1.82	0.84	0.06		

表 8: ”love”

	N	Mean	Std.Deviation	Std.Error	F	Sig.
0	9042	3.97	1.36	0.01	553.465	0.000
1	2428	4.65	0.85	0.02		

表 9: ”great”

	N	Mean	Std.Deviation	Std.Error	F	Sig.
0	8700	3.99	1.37	0.02	346.118	0.000
1	2770	4.51	0.93	0.02		

这表明含不含关键词对星级的分布有显著影响。

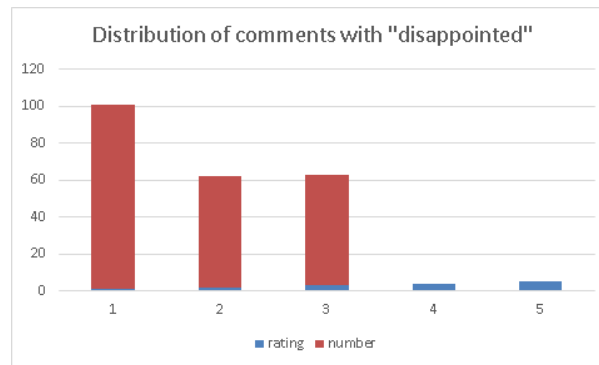


图 1: The parts of plastics' life cycle that might cause damage to the environment.

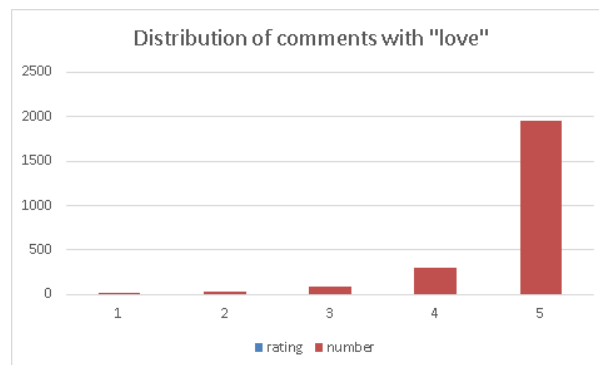


图 2: The parts of plastics' life cycle that might cause damage to the environment.

下面我们用柱状图表示具体的分布:

从图中我们可以得出以下结论:

- 含关键词“disappointed”的评论和星级 1 密切相关;
- 含关键词“love”或“great”或“good”的评论和星级 5 密切相关。

Maximal plastic waste is a problem related to plastic production, using, and recycling. Based on LCA, we analyzed all the pollution generated in the life cycle of plastics, and then reduced this

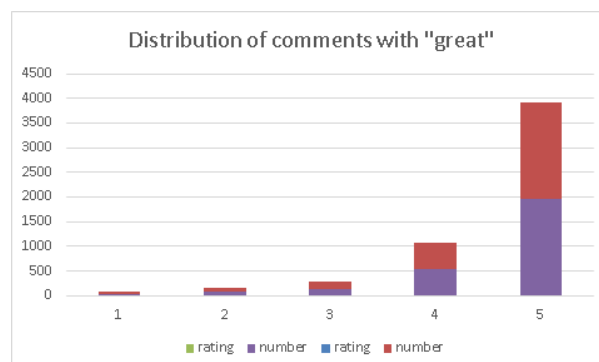


图 3: The parts of plastics' life cycle that might cause damage to the environment.

problem into a linear programming model. Finally, we use the simplex method to solve based on an example.

5.6 Theories related to our model

In this section, we will introduce theories that are closely related to our model.

5.6.1 PLC theory

To guarantee that the maximal usage or producing of plastics will not damage the environment further, we identify the parts of plastics' life cycle that might cause damage to the environment, as is shown in figure 4. Consider a revised product life cycle (PLC) theory: as every product has its life cycle from R&D(earch and development) to launching on market to decline and exit the market as an end, plastics also cycle from products to waste. For plastic products from various source, their producing processed may affect the environment as well as their managing courses. The model expand based on the plastic life cycle theory.

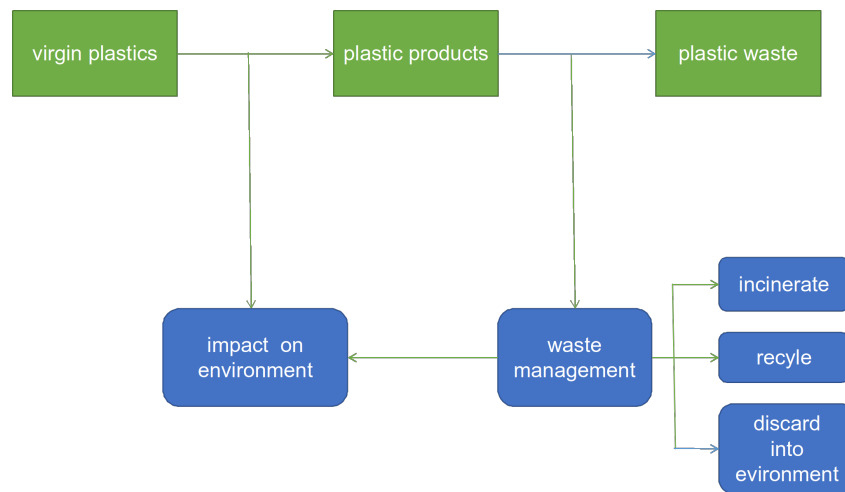


图 4: The parts of plastics' life cycle that might cause damage to the environment.

5.6.2 Linear programming

Linear programming[18] (LP) is a tool for solving optimization problems for which we do following:

1. We attempt to maximize (or minimize) a linear function of the decision variables which is called the objective function.
2. The values of the decision variables must satisfy a set of constraints. Each constraint must be a linear equation or linear inequality.
3. A sign restriction is associated with each variable. For any variable x_i , the sign restriction specifies that x_i must be either nonnegative ($x_i \geq 0$) or unrestricted in sign.

So we try to set our objective function as the permitted maximal amount of plastic and abstract the most limited conditions into constraints.

5.7 Qualify the impact on the environment

We analyzed the pollution caused by each stage of the plastic's entire life cycle.

5.7.1 Producing process

According to the first assumption, we consider the annual production and waste as the same variable S . And naturally it is divided into various sorts of plastic.

$$S = \sum_i s_i \quad (3)$$

Note that, i represents a sort of plastic and s_i is the annual production of this sort of plastic. All sorts of plastic taken into account are PVC, PE, PP, PS, PET, PUR and PC[11].

To quantify to what extent plastic waste impact the environment in its life cycle, we use Emissions air(EA) and Emissions water (EW) to measure emissions, whose units are Units of Polluted Water (UPW /ton) and Unit of Polluted Air with the production of a ton of plastic(UPA /ton). UPW is the number of cubic meters of water polluted up to the European drinking water standard by the production of 1 tonne of the material. UPA is a similar measure for air and gives the cubic meters of air needed to dilute the emissions to the European maximum acceptable concentration (MAC)[11]. Then the Emissions of water and air can be formulated below:

$$TEA^\alpha = \sum_i s_i \times EA_i^\alpha \quad (4)$$

$$TEW^\alpha = \sum_i s_i \times EW_i^\alpha \quad (5)$$

Given the energy consumption per unit of plastic production, we set energy consumption of specific plastic production as E_i and the proportional coefficient of carbon emission per unit of energy consumption as k , total emission of CO2 can be calculated as:

$$TEC^\alpha = k \sum_i s_i \times E_i^\alpha \quad (6)$$

5.7.2 Managing process

When it comes to plastic waste management, there are three possible fates for managed plastic waste: being recycled, being incinerated and being discarded (be landfilled or be discarded in natural environment). According to the third and fourth assumption, we assume there are only two managing method: incinerating or recycling.

For the proportion of recycled PW, they are recycled by The material recycling facilities(MRFs). The mechanical separation removes polyethylene terephthalate (PET) and high density polyethylene(HDPE) bottles at high efficiency, which are sent to recycling, and other minor high calorific

fluxes sent to energy recovery in cement kilns[4]. After that we could get certain recovery rate, we set this rate as $1 - \sigma$, the proportion of incinerated part is σ .

Firstly, we apply a classical environmental assessment method: life cycle assessment (LCA), to evaluate the environmental issues associated with solid waste with solid waste management[5].

Secondly, we simplify and abstract the many categories of LCA into three categories based on data availability and computational matching: EA, EW and EC.

Furthermore, because the recycling technology(MRFs) saves this part of PW from incineration and put recycled products into market again, while the harmful emissions from recycling are often less than the emissions from incinerating this part of the plastic[15]. There are digital supports that reveal the positive influence about of recovery, as is shown in figure 5 and figure 6. So we calculate this different value as compensatory emissions.

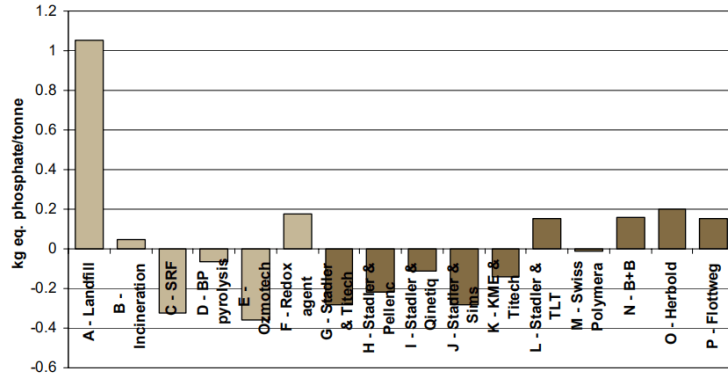


图 5: Net eutrophication potential.

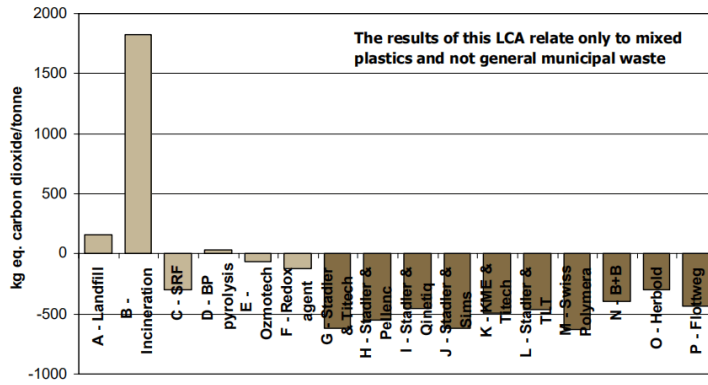


图 6: Net global warming potential.

Then the environmental impact of plastic waste management can be calculated:

$$TEA^{\beta} = \sigma \sum_i s_i \times EA_i^{\beta} - (1 - \sigma) \sum_i s_i \times CEA_i^{\beta} \quad (7)$$

$$TEW^\beta = \sigma \sum_i s_i \times EW_i^\beta - (1 - \sigma) \sum_i s_i \times CEW_i^\beta \quad (8)$$

Apart from above analysis, although we assume that all the plastic waste discarded in natural environment will be managed, too much discarded plastic waste (even if temporarily) will notably do harm to marine ecosystem[3]. Thus it's necessary to estimate the mass of plastic waste(symbolized by SM) inputs from land into the ocean, given specific total plastic waste.

$$SM = \varsigma \sum_i s_i \quad (9)$$

For the plastic waste that has been incinerated, in the process of incineration, we take the emission of greenhouse gases CO_2 (EC) to represent it's contributions to global warming:

$$TEC^\beta = \sigma \sum_i s_i \times N_i - \mu(1 - \sigma) \sum_i s_i \quad (10)$$

5.8 Constraints given by environmental carrying capacity

Firstly, according to the fifth assumption, the pollutant emission constraints are roughly estimated by the percentage of GDP plastic industry contributes, quantified by coefficient θ .

$$TEA^\alpha + TEA^\beta \leq \theta V_{air} \quad (11)$$

$$TEW^\alpha + TEW^\beta \leq \theta V_{water} \quad (12)$$

Secondly, when it comes to greenhouse gas CO_2 , it's more propriate to use its fossil fuel consumption percentage to set the restraint because the fossil fuel combustion is the principal factor of the global warming, rather than natural carbon cycles[11].

$$TEC^\alpha + TEC^\beta \leq \iota V_{CO_2} \quad (13)$$

Last but not least, considering current quantity of plastic discarded in the ocean don't exert severe impact on marine ecosystem, we regard zero net input into the sea an acceptable circumstance for marine environment, i.e. annual plastic waste input can't beyond the quantity of annual management of marine plastic(symbolized by M).

$$SM \leq M \quad (14)$$

5.9 Model paradigm

In the end, our model can be written as follows:

Objective functions:

$$S = \sum_i s_i \quad (15)$$

subject to:

$$\begin{cases} TEA^\alpha + TEA^\beta \leq \theta V_{air} \\ TEW^\alpha + TEW^\beta \leq \theta V_{water} \\ TEC^\alpha + TEC^\beta \leq \iota V_{CO_2} \\ SM \leq M \end{cases} \quad (16)$$

5.10 Analysis of China using our model

In this section, we will combine the data from China to further introduce the model.

China is a developing country which emission largest amount of PW to marine, which may be sensitive to change of plastic amount. As shown in figure 7, China has the most and most developed plastic industry in the world. We apply our model to the practical examples of China. Table 10 lists seven disposable plastics mainly used in China. We consider them as model variables s_i .



图 7: China's developed plastic industry.

We used data from the National Bureau of Statistics of China[16] and obtained parameters closely related to the model. Specific information can be consulted in the appendix.

Note that for practical reasons, we set some settings for the model, where the composite emission to water and air of recycling process should be deemed as zero for the following considerations:

表 10: The main types of disposable plastics used in China

Variable	Plastic types
s1	PVC
s2	PE
s3	PP
s4	PS
s5	PET
s6	PUR
s7	PC

1. The proportion of recycled PW is relatively small in realistic practices.
2. Most regions of China are not developed enough to have the advanced technology of MRFs to manage PW.
3. Chinese PW sorting and managing system are not perfect.

Finally, the linear programming model based on Chinese data is as Eq.17 and Eq.18:

$$S = s_1 + s_2 + s_3 + s_4 + s_5 + s_6 + s_7 \quad (17)$$

$$\begin{cases} 700s_1 + 265s_2 + 325s_3 + 255s_4 + 180s_5 + 0s_6 + 180s_7 \leq 2.88 \times 10^{11} \\ 3000s_1 + 1650s_2 + 3685s_3 + 6355s_4 + 8000s_5 + 0s_6 + 5050s_7 \leq 7.89696 \times 10^{10} \\ 1.99s_1 + 4.42s_2 + 4.43s_3 + 4.77s_4 + 3.23s_5 + 0.49s_6 + 1.96s_7 \leq 8.062 \times 10^{11} \\ 0.41 \times (s_1 + s_2 + s_3 + s_4 + s_5 + s_6 + s_7) \leq 4.85 \times 10^5 \end{cases} \quad (18)$$

We solved this problem by using the simplex method through MATLAB, and obtained the largest amount of plastic that China can bear in one year. The results show that China can bear 1.18 million tons of plastic a year. When the optimal solution is obtained, the production volume of various plastics is shown in table11.

We will continue to analyze this case later to explore the impact of different constraints.

6 Plastic waste reduced

At present, the global use of plastics has exceeded the upper limit of environmental tolerance, which poses a huge threat to the ecological environment. Considering the huge differences in the existence of countries around the world, we must consider each region separately. In section 6, we propose a model of Happiness based on SVR[6], which we call HSVR. Using HSVR, we will explore the impact of the use of plastic on human living standards in a region, and try to reduce the amount of plastic used as much as possible without affecting much of human living standards.

表 11: The amount of various plastics in the optimal solution

Plastic types	Amount of plastic(t)
PVC	210428.703365915
PE	107339.181740866
PP	174475.173021043
PS	150364.822097472
PET	233445.923823156
PUR	42560.4917899467
PC	264312.533429888

Our method can be summarized as follows: the use of plastic is essential, and reducing the use of plastic will inevitably affect the living standards of local people. By analyzing the characteristics of different regions (especially the amount of plastics consumed) and the people's living standards, we find out the relationship between the amount of plastics consumed and the people's happiness index. Based on this, we can analyze how much plastic consumption we can reduce without significantly reducing the happiness index.

The following of this section will first briefly introduce the method we use, and then obtain a specific regression model based on the Global Happiness Report[7] and plastic consumption in each region[8].

6.1 SVR algorithm

Support vector regression(SVR) is an application of SVM (support vector machine) to regression problems. Support vector machines construct a hyperplane or a series of hyperplanes in a high-dimensional or infinite-dimensional space, which can be used for classification, regression, or other tasks. Intuitively, using a hyperplane to achieve a good segmentation can make the closest training data points have the largest separation distance in any category. This is because usually a larger margin can have a lower generalization error of classifier[9].

Given training vectors $x_i \in R^p, i = 1, \dots, n$, and a vector $y \in R^n$. ε -SVR solves the following primal problem:

$$\min_{\omega, b, \zeta, \zeta^*} \frac{1}{2} \omega^T \omega + C \sum_{i=1}^n (\zeta_i + \zeta_i^*) \quad (19)$$

subject to

$$y_i - \omega^T \phi(x_i) - b \leq \varepsilon + \zeta_i \quad (20)$$

$$\omega^T \phi(x_i) + b - y_i \leq \varepsilon + \zeta_i^* \quad (21)$$

where $\zeta_i, \zeta_i^* \geq 0, i = 1, \dots, n$.

Its dual is

$$\min_{\alpha, \alpha^*} \frac{1}{2}(\alpha - \alpha^*) + \varepsilon e^T(\alpha + \alpha^*) - y^T(\alpha - \alpha^*) \quad (22)$$

subject to

$$e^T(\alpha - \alpha^*) = 0 \quad (23)$$

where $0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, \dots, n$, e is the vector of all ones, $C > 0$ is the upper bound, Q is an n by n positive semidefinite matrix, $Q_{ij} \equiv K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ is the kernel. Here training vectors are implicitly mapped into a higher (maybe infinite) dimensional space by the function ϕ . ϕ can be linear, polynomial, sigmoid or others.

The decision function is:

$$\sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x_i, x) + \rho \quad (24)$$

For more details, please refer to [6].

6.2 Detailed analysis

In this section, we will introduce the quantitative standard of happiness, analyze this problem, and use SVR to get the function of happiness and plastic consumption \hat{f} so that $\hat{f} : X \mapsto y$, where X is a vector in $\mathcal{X} = \mathcal{R}^d$, including a dimension of plastic consumption.

6.2.1 The quantitative standard of happiness

We use the happiness quantification standard used by the World Happiness Report[7], where the happiness scores and rankings use data from the Gallup World Poll. The scores are based on answers to the main life evaluation question asked in the poll. This question, known as the Cantril ladder, asks respondents to think of a ladder with the best possible life for them being a 10 and the worst possible life being a 0 and to rate their own current lives on that scale. The scores are from nationally representative samples for the years 2013-2016 and use the Gallup weights to make the estimates representative.

Obviously, people's happiness depends not only on the amount of plastic consumed. In fact, plastic consumption and waste account for only a small part of the factors affecting people's happiness. There are six factors usually concerned to be relevant to happiness: economic production, social support, life expectancy, freedom, absence of corruption, and generosity. The data that estimates the extent to which each of six factors contribute to making life evaluations higher in each country than they are in Dystopia, a hypothetical country that has values equal to the world's lowest national averages for each of the six factors, is also used to as part of the input of HSVR, with some preprocessing:

$$x^* = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (25)$$

where x_{max}, x_{min} is the maximum and minimum of the original data, respectively.

Note that the first preprocessing maps the data to $[0, 1]$. The preprocessing step will erase the differences between the data formats and make the characteristics of the data more obvious.

6.2.2 Plastic Emphasized

The impact of the consumption and use of plastic on happiness should be far less than that of the other six factors mentioned above, which means that the input to the problem should be a weighting of all seven factors:

$$X = (\epsilon_1 x_1, \epsilon_2 x_2, \epsilon_3 x_3, \epsilon_4 x_4, \epsilon_5 x_5, \epsilon_6 x_6, \epsilon_7 x_7) \quad (26)$$

However, what we need to find is mainly the relationship between plastic consumption and happiness, so we give plastic consumption a sufficiently high weight, where $\epsilon_i = 1, i = 1, \dots, n$.

The lack of data[10] in research on plastic production and consumption has always been a serious problem. To the best of our knowledge, there are no specific statistics on the amount of plastic used in many regions, especially in developing countries. This means that our research will face problems of insufficient data volume and potential data imbalances. Based on this, we chose SVR as our classifier, because SVR has the following characteristics[9] and is suitable for solving this problem:

- Effective in high dimensional spaces.
- Versatile: different Kernel functions can be specified for the decision function.

6.3 Model parameter selection

We randomly selected 20 countries from [7] and [8] as our training set as our training set. These countries include countries that consume more plastic, such as China and India, and developed countries, such as the United States and the United Kingdom. In combination with the training situation, we adjusted the model in real time to obtain the optimal model.

6.3.1 Model performance evaluation

We choose the Euclidean norm of the regression value and the true value(or the mean square error, MSE) to measure the quality of the regression result:

$$MSE = \frac{1}{k} \|y - \hat{y}\|_2 \quad (27)$$

where y is the true value, and \hat{y} is the regression value.

6.4 Specific model parameters

We have implemented four kinds of kernel function[9] of SVR and verified them separately.

The four cores are:

- linear: $\langle x, x' \rangle$
- polynomial: $(y\langle x, x' \rangle + r)^d$.
- rbf: $e^{-\gamma\|x-x'\|^2}$.
- sigmoid: $\tanh(\langle x, x' \rangle + r)$.

Due to the high feature dimension of our training set and the small number of training samples, in general, the effect of the linear kernel should be better. The results in figure 8 also illustrate this:

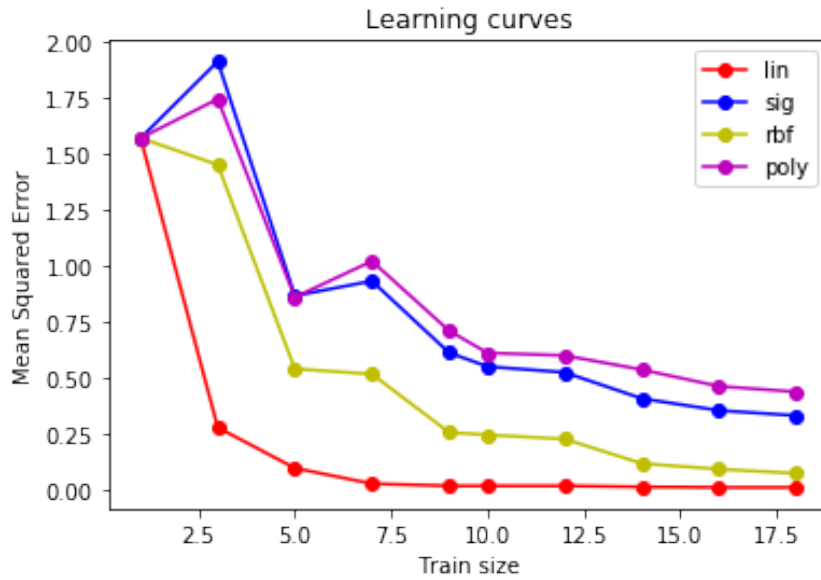


图 8: Training results of different SVR kernel functions.

As can be seen from figure8, the results of linear and rbf are much better than the results of sigmoid and polynomial, which all converge to an error close to 0. In fact, the training error of the linear kernel is on the order of 10^{-2} . Moreover, the linear kernel has already converged by training about ten samples, which shows that our data volume is sufficient for our model.

6.5 Predicting minimum plastic usage

From the above analysis, we find that the linear kernel function has the best fitting effect, indicating that the plastic consumption has a linear relationship with the living standard. This is consistent with per capita plastic consumption in different regions. Table 12 is a table of plastic consumption in different regions. Intuitively, the consumption of plastic is positively related to the

表 12: Plastic consumption in different regions

Area	Plastic consumption (1 ton per 100,000 people)
China	1316.1
North America	1571.9
Asia Pacific	657.1
Western Europe	2870.6
India	251.9
Middle East	580.4
Central and South America	308.8
Central and Eastern Europe	1468.4
Africa	99.8
Japan	812.9

living standard of the region. Developed regions generally have higher plastic consumption. Of course, this is also related to other factors. These are reflected in our model.

Based on this conclusion, we found that reducing the amount of plastic used will inevitably affect people's happiness. Therefore, we need to make a compromise between happiness and environment. This is undoubtedly a very painful thing. Therefore, we need to find a threshold on the tolerance of the regions κ . κ is such a value: in a specific area, if happiness $\Gamma = \hat{f}(X) < \kappa$, it is not worthwhile to reduce consumption. Note that \hat{f} is trained in section 6.3. κ can be derived from historical data and is closely related to each place.

Unfortunately, many people are unwilling to pay a lot for environmental protection[14], as is shown in figure 9. So we carefully set this value to 0.5% of the existing happiness index.

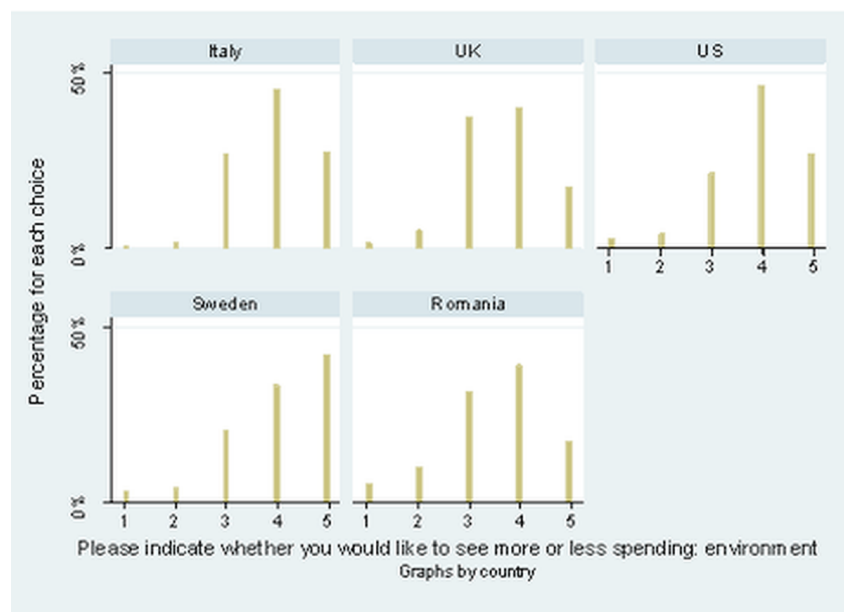


图 9: Happiness-Plastic curve of US.

For example, figure 10 shows this relationship in the US. Currently, plastic consumption in the United States is 15.719 kg per people. If the amount of plastic used per person per year in the United States is reduced to 6.19 kg per year, the happiness level in the United States will be reduced by 0.5%. We think of this as the largest reduction in plastic use in the United States.

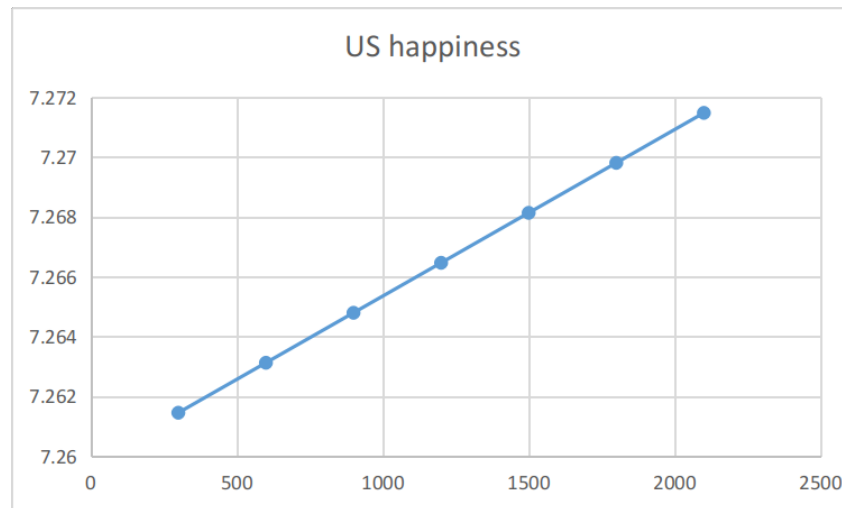


图 10: Happiness-Plastic curve of US.

7 The minimal target of PW and its impacts

Given that we have get the defined maximal amount of palstic production and the most extent that the plastic waste can be reduced to, taking a statistics of all the regions in the world, we can then get the lowest global target. Next, we analyzed the impact of this measure on human life, environment and plastic industry.

7.1 Basic Theories

We have used some theories in this section as a guide, which will be briefly introduced below.

7.1.1 EASEWASTE

TheEnvironmental assessment of solid waste systems and technologies (EASEWASTE) is a model that enables waste managers to assess environmental impacts of solid waste management systems and to make comparisons among recovery technologies for a region with a given population and waste generation. Users can define all necessary data for waste composition, collection, treatment, recovery and disposal and life cycle inventory data for materials and energy used in the waste management system[5].

7.1.2 CGE

The CGE model explicitly assumes that the behavior of all economic subjects is optimized in accordance with the usual neoclassical microeconomic theory, so it is a model concerned with the behavior of general rather than local economic subjects. It uses the assumption that markets are balanced, not unbalanced, and that all markets are settled at the same time. The basic economic units it analyzes are producers, consumers, governments and foreign economies[20].

7.1.3 AHP

AHP has particular application in group decision making and is used around the world in a wide variety of decision situations, in fields such as government, business, industry, healthcare, shipbuilding and education.

Rather than prescribing a "correct" decision, the AHP helps decision makers find one that best suits their goal and their understanding of the problem. It provides a comprehensive and rational framework for structuring a decision problem, for representing and quantifying its elements, for relating those elements to overall goals, and for evaluating alternative solutions[21].

7.2 Global Target

The global goal should be the sum of the small goals in each region. So we divided the world into 10 typical countries according to their plastic assumption amount, as is shown in table 13. In a certain category, there are similar plastic assumptions and culture background in a certain category.

For each region, we used the model explained in Section 6 to calculate the target for minimum plastic use. Global goals can be achieved by adding them together. All details are shown in table 13.

Note that the total achievable minimal plastic consumption obtained is about 6.6 hundred million tons.

7.3 Impact Path Analysis

In this section we discuss three most relevant aspects that can be changed: human life, environment and plastic industry.

表 13: Estimated minimum usage for every region.

Region	Polulation	Minimum plastic	Estimated minimum usage
China	13.86	425	5890.42
North America	3.70	619	2290.49
Asia Pacific	7.12	108	756.01
Western Europe	1.41	1975	2784.72
India	13.39	2700	37087.65
Middle East	4.90	1875	9188.28
Central and South America	6.30	1050	6614.18
Central and Eastern Europe	1.21	620	744.01
Africa	12.16	49	596.05
Japan	1.28	230	291.62

7.3.1 How human life will be alert

In section 6, we have discussed how people's happiness index will be affected by the reduction of plastic usage. But happiness only reflect people's satisfaction and quality of life, there are still other possible changes of human life.

For example, living style of most people may be changed. Because the supply of plastic is less than the demand, the price of plastic products is bound to rise. People will be pushed to choose products made of other materials as the succedaneum, which may lead to popularity of new materials in the market.

But in the long run, less pollution and more substantial development will lead a better quality of life to every one.

7.3.2 How environment will be affected

It should be noticed that there is a 2.2 hundred million difference between the maximal and minimal amount of plastics. Since there are part of plastic plastics that should be recycled for realizing the target reduction. Based on LCF thoery, we can use EASEWASTE theory to estimate the positive externalities on the environment of the recovery. The environmental impacts are calculated based on the overall environmental exchanges which comprise resource consumptions and emissions to air, water and soil which in the life cycle impact assessment may contribute to environmental impact potentials[5].

In section 5, we have quantified the negative impacts of plastic production or plastic waste under quite a few assumptions. When the plastic waste levels are reduced to target values, the emission harmful gases and carbon dioxide will decline proportionally. The pollution caused by this 2.2 hundred million plastic waste will also be migrated.

7.3.3 How to influence plastic industry

This is one of the few negative effects of this achievement. Some skeptics insist that once limite the production of plastics, the muti- trillion-dollar plastic industry will significantly shrink, which is bad for economic balance especially for countries that developed on primary and secondary industries like China. The accounting loss of a certain country or region could be calculated as:

$$L = O \times P \quad (28)$$

where L refers to loss, O is annual output of plastics, and P is annual industry average price of plastics.

Additionally, we should also consider the economic loss or benefit which enrolled the opportunity cost. Computable General Euilibrium(CGE) provide us a fair method to analysis the general influence of this reduction to the plastic industry or even the whole economic.

CGE suggest that when plastic market come to an equilibrium again after the reduction:

1. Producer will switch to other industries and may lead to prosperous of some new and clean

materials, which will improve economic growth and productivity release.

2. Consumers will be willing to change their consuming habits and more likely to accept new products, which will boost domestic demand and stimulate consumption.
3. The government can put up new policies to reduce the plastic production by adjusting taxes and import and export quotas, which may adjust the tax and tariff policy and make more fiscal revenue.

7.3.4 Total impact

In general, the total impact should be a linear combination of the effects of each factor:

$$I = \sigma_1 i_1 + \sigma_2 i_2 + \sigma_3 i_3 \quad (29)$$

where σ_1 , σ_2 and σ_3 are the linear coefficients of different factors.

AHP is used to list the matrix and determine the weight of each factor on the whole society. Since the above three factors are abstract social phenomena, we can first quantify the impact with some indicators. Since quantized results have different orders of magnitude and units, we normalize these effects. Multiply the result by the weight of each factor to get the total impact.

Using AHP, we can determine the value of σ_1 , σ_2 and σ_3 so that we can calculate the overall impact of different regions. Generally speaking, the weight of different influencing factors in different regions should be different.

8 Discussion about equity

The plastic problem is a global problem and requires international cooperation to resolve it. However, the situation is different in different countries or regions, and we cannot generalize. Different countries have different industrial structures and are affected to varying degrees by the reduction in plastic use. Intuitively, the more a country is affected by plastic, the more we should tolerate this country and allow them to reduce the use of plastic at a slower rate. Therefore, in this section, we further analyze the model proposed in section 5, and give specific methods to analyze the factors most affected by plastics in a region. Based on this method and continued analysis on Chinese cases, we further give theoretical guidance and relevant suggestions.

8.1 Shadow price

We use shadow prices to evaluate the sensitivity of linear programming models. A shadow price is commonly referred to as a monetary value assigned to currently unknowable or difficult-to-calculate costs. It is based on the willingness to pay principle—in the absence of market prices, the most accurate measure of the value of a good or service is what people are willing to give up in order to get it[17]. Shadow price reflects the value of the optimal use of resources. In the model in Section 5, we measured four resources: water quality, air quality, carbon emissions, and marine debris recycling. The shadow prices of these four resources reflect the impact of the region's reduction in plastics on these four resources.

In our model, resources fall into two categories:

- Natural resources. This type of resource is determined by the region itself, and we cannot change the total amount of this resource in a certain region, such as air and water. If the shadow price of this type of resource is high, then we have no choice but to tolerate this area, because forcing such areas to reduce the production and use of plastic will inevitably cause a great burden on the local area.
- Controllable resources. This is a resource that can be determined by the actions of a country, such as carbon emissions and the amount of marine debris recovered. If the shadow price of this type of resource is high, we must ask the country or region because it is the inaction of this country that has caused the pollution of plastic waste.

For plastic production in a country or region, empirical conclusions often look like this: Developed countries often have better management of plastics and should pay more attention to the issue of natural resources. Developing countries need a lot of resources for development and they should be required to strengthen their management of plastics.

However, this conclusion is not widely applicable, and there are always exceptions. Therefore, for a specific region, the best way is to calculate the shadow price of each resource, so as to truly determine the impact of each resource.

8.2 Determining shadow prices using dual theory of linear programming

Two dual linear programming problems can be described like this[18]:

$$\max z = CX \quad (30)$$

subject to:

$$\begin{cases} AX \leq b \\ X \geq 0 \end{cases} \quad (31)$$

and

$$\max \omega = b^T Y \quad (32)$$

subject to:

$$\begin{cases} A^T Y \geq C^T \\ Y \geq 0 \end{cases} \quad (33)$$

where $C = (c_1, c_2, \dots, c_n)$, $X = (x_1, x_2, \dots, x_n)^T$, $y = (y_1, y_2, \dots, y_m)^T$, A is a matrix with m rows and n columns, $b = (b_1, b_2, \dots, b_m)^T$.

The shadow price of resources is the dual solution of the linear programming problem Y^* for the following properties:

- Weak duality. For any feasible solution of the original problem X and any feasible solution of the dual problem Y , there are $CX \leq b^T Y$.
- Complementary relaxation. If $y_i^* > 0$, then $\sum_{j=1} na_{ij}x_j^* = b_i$; If $\sum_{j=1} na_{ij}x_j^* < b_i$, then $y_i^* = 0$.

Based on this conclusion, we can get the following inference: while obtaining the maximum value of the objective function, the Lagrange multiplier[19] corresponding to each constraint is the shadow price of the resource corresponding to this constraint. Therefore, we only need the Lagrange multiplier for the optimal solution to obtain the shadow price.

8.3 Shadow price of China's plastic constraint equation: an example

In this section, we further analyze the constraint equation of plastic consumption in China Eq.17 and Eq.18. Based on this, we give the best way for international cooperation for China.

We use MATLAB to obtain the Lagrange multiplier of Eq.17 and Eq.18 when the optimal solution is obtained, as shown in table 14.

As can be seen from the table 14, the shadow price of China's marine garbage recycling volume is much higher than the shadow prices of other influencing factors. This means that the amount of plastic that China can use is actually largely constrained by the amount of marine debris recovered. Part of China's plastic has entered the ocean, but not a lot of it has been recycled. This is an urgent problem. In fact, our model shows that once the constraint is removed, the amount of plastic that can be used normally in China will be much larger than the existing more than one hundred piles, reaching tens of millions of tons. Such a reduction is actually an acceptable burden for China as is analysed in section 6.

In terms of natural resources, air has the largest shadow price. For China, priority should be given to reducing the production and use of plastics with more air pollution as much as possible, which will have the most obvious ecological benefits to the region.

9 Strength and weakness

The strengths and Weaknesses of our model are summarized as follows:

表 14: Shadow prices of Chinese resources

Limiting factor	Shadow price
Air	$1.63233961552303 \times 10^{-14}$
Water	$1.12005836576517 \times 10^{-17}$
Carbon emission	$5.66607374733818 \times 10^{-16}$
Marine garbage recycling	2.43902439024392

9.1 Strengths

Model complexity. Our model incorporates relevant research results and considers many details, making the model complicated enough to include enough factors. Therefore, our model can analyze the situation in specific regions and draw targeted conclusions, which is conducive to practical use.

Agreement between experimental data and real data. In the process of estimating maximal plastic waste, our model based on PLC theory carefully consider the environmental impact of plastic in different phases of its life cycle rather than merely in the phase of management. What's more, despite the lack of data, our machine learning model still fits the data well. The results of tests on real data sets also validate the performance of our model, where experimental data and real data agrees well.

The amount of data. In the process of modeling, we extensively collected relevant industry data, making the model sufficiently universal and extensible. To our knowledge, few studies on the relationship between plastics and the environment have been able to model using so much data.

9.2 Weaknesses

Data accuracy of the model. In order to consider as many factors as possible, our model introduces more parameters. This makes it difficult to obtain our data, especially when related research is very scarce. Further we will try to simplify the model for different scenarios and improve the usability of the model.

Macro analysis. Our model is too detail-oriented and has insufficient performance in macro analysis. In global research, our model encountered some difficulties and showed some uncertainty. In the future, we will try to explore the relationship between the overall situation and the details, and strive to let our model to show the superiority in global problems, too.

10 Conclusion

To mitigate the plastic waste problem, we establish a linear programming base model to estimate the environmental impacts of plastic waste, and get the maximal plastic waste amount without further damage to environment. It is proved that the most extent plastic can be reduced to could be confirmed by HSVR model. A minimal target amount of plastic usage is achieved by expending HSVR model. The equity issue can also be solved by dual programming and sensitivity analysis of factors. Finally, we reckon a timeline to approach this realistic minimal target of plastic using amount.

参考文献

- [1] Geyer, Roland, Jenna R. Jambeck, and Kara Lavender Law. "Production, use, and fate of all plastics ever made." *Science advances* 3.7 (2017): e1700782.
- [2] Giacobelli, Claudia. "Single-Use Plastics: A Roadmap for Sustainability." (2018).
- [3] LI, Wai Chin, H. F. Tse, and Lincoln Fok. "Plastic waste in the marine environment: A review of sources, occurrence and effects." *Science of the Total Environment* 566 (2016): 333-349.
- [4] Rigamonti, Lucia, et al. "Environmental evaluation of plastic waste management scenarios." *Resources, Conservation and Recycling* 85 (2014): 42-53.
- [5] Kirkeby, Janus T., et al. "Environmental assessment of solid waste systems and technologies: EASEWASTE." *Waste Management & Research* 24.1 (2006): 3-15.
- [6] Smola, Alex J., and Bernhard Schölkopf. "A tutorial on support vector regression." *Statistics and computing* 14.3 (2004): 199-222.
- [7] World Happiness Report. <https://kaggle.com/unsdsn/world-happiness>. Accessed 15 Feb. 2020.
- [8] "Plastic Industry Worldwide." Statista, <https://0-www-statista-com.lib.rivier.edu/study/51465/global-plastics-industry/>. Accessed 15 Feb. 2020.
- [9] 1.4. Support Vector Machines —Scikit-Learn 0.22.1 Documentation. <https://scikit-learn.org/stable/modules/svm.html>. Accessed 15 Feb. 2020.
- [10] Groot, Jim, et al. "A comprehensive waste collection cost model applied to post-consumer plastic packaging waste." *Resources, Conservation and Recycling* 85 (2014): 79-87.
- [11] Andrady, Anthony L., ed. *Plastics and the Environment*. John Wiley & Sons, 2003.
- [12] "Cross-Validation (Statistics)." Wikipedia, 7 Feb. 2020. Wikipedia, [https://en.wikipedia.org/w/index.php?title=Cross-validation_\(statistics\)&oldid=939673086](https://en.wikipedia.org/w/index.php?title=Cross-validation_(statistics)&oldid=939673086).
- [13] McLachlan, Geoffrey J., Kim-Anh Do, and Christophe Ambroise. *Analyzing microarray gene expression data*. Vol. 422. John Wiley & Sons, 2005.
- [14] Arpad, Todor. "Willing to Pay to Save the Planet? Evaluating Support for Increased Spending on Sustainable Development and Environmentally Friendly Policies in Five Countries." *PLOS ONE*, vol. 13, no. 11, Nov. 2018, p. e0207862. PLoS Journals, doi:10.1371/journal.pone.0207862.
- [15] Shonfield, Peter. "LCA of management options for mixed waste plastics." WRAP, UK (2008).
- [16] Country data. <http://data.stats.gov.cn/>. Accessed 17 Feb. 2020.
- [17] "Shadow Price." Wikipedia, 11 Dec. 2019. Wikipedia, https://en.wikipedia.org/w/index.php?title=Shadow_price.
- [18] Hillier, Frederick S. *Introduction to operations research*. Tata McGraw-Hill Education, 2012.

-
- [19] “Lagrange Multiplier.” Wikipedia, 17 Dec. 2019. Wikipedia, https://en.wikipedia.org/w/index.php?title=Lagrange_multiplier&oldid=931158065.
- [20] CGE model-encyclopedia of MBA think tanks. <https://wiki.mbalib.com/wiki/CGE>. Accessed 18 Feb. 2020.
- [21] Analytic Hierarchy Process - Wikipedia. https://en.wikipedia.org/wiki/Analytic_hierarchy_process. Accessed 18 Feb. 2020.