

Evaluation of system efficiency using the Monte Carlo DEA: The case of small health areas

使用蒙特卡罗DEA评价系统效率：以医疗领域为例

European Journal of Operational Research, 2014

Mercedes Torres-Jiménez^{a,*}, Carlos R. García-Alonso^b, Luis Salvador-Carulla^c,
Vicente Fernández-Rodríguez^d

^a Department of Mathematics and Engineering, Universidad Loyola Andalucia

^b Department of Mathematics and Engineering, Universidad Loyola Andalucia, Escritor Castilla Aguayo, 4. 14004 Córdoba, Spain

^c Brain and Mind Research Institute, and Centre for Disability Research and Policy Faculty of Health Sciences, University of Sydney, Sydney, Australia

^d Department of Business Organization, Universidad Loyola Andalucia

目录

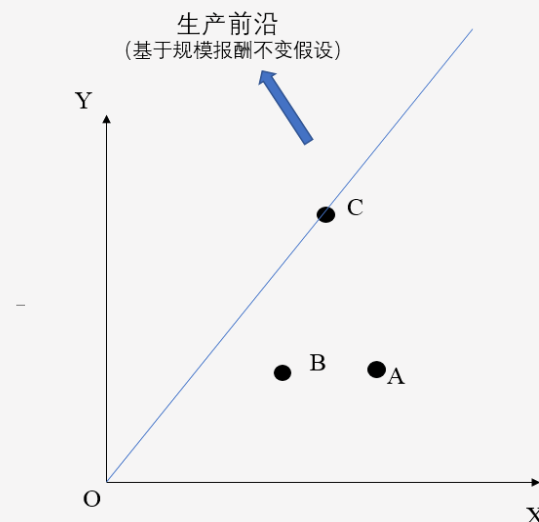
1. 问题来源
2. 研究内容
3. 问题解决效果、创新点
4. 启发

1 问题来源

效率：描述资源使用的情况，衡量投入转换为产出的过程

效率评价对象，称为**决策单元** (DMU, Decision Making Unit,)

- 机场：员工、成本；旅客人数、收入
- 银行：存款、员工；放款、投资
- 大学：教职人数、资产；学生数量、收入
- 区域医疗系统



数据包络分析 (DEA, Data Envelopment Analysis) 是一种效率评价方法，其本质是线性规划。

以 **(权重*投入指标) / (权重*产出指标)** 作为效率，总投入最小、总产出最大的决策单元构成生产前沿。

1 问题来源

$$\begin{aligned} \max \quad & \theta_o = \frac{\sum_{r=1}^s \mu_r y_{ro}}{\sum_{i=1}^m v_i x_{io}} \\ s.t \quad & \sum_{r=1}^s \mu_r y_{rj} - \sum_{i=1}^m v_i x_{ij} \leq 0 \quad (j=1, 2, \dots, n) \\ & v_i \geq 0 \quad (i=1, 2, \dots, m) \\ & \mu_r \geq 0 \quad (r=1, 2, \dots, s) \end{aligned}$$

对每一个决策单元，分别构建线性规划

已知：每个决策单元的投入产出指标值

1.变量：**投入权重**；**产出权重**。

2.目标函数：最大化效率=

(投入权重*投入指标) / (产出权重*产出指标)

3.约束：**权重的合理性**，即这组权重用于其他决策单元时，计算的效率在[0,1]之间；权重非负。

基础模型在复杂系统中应用时，存在不足：

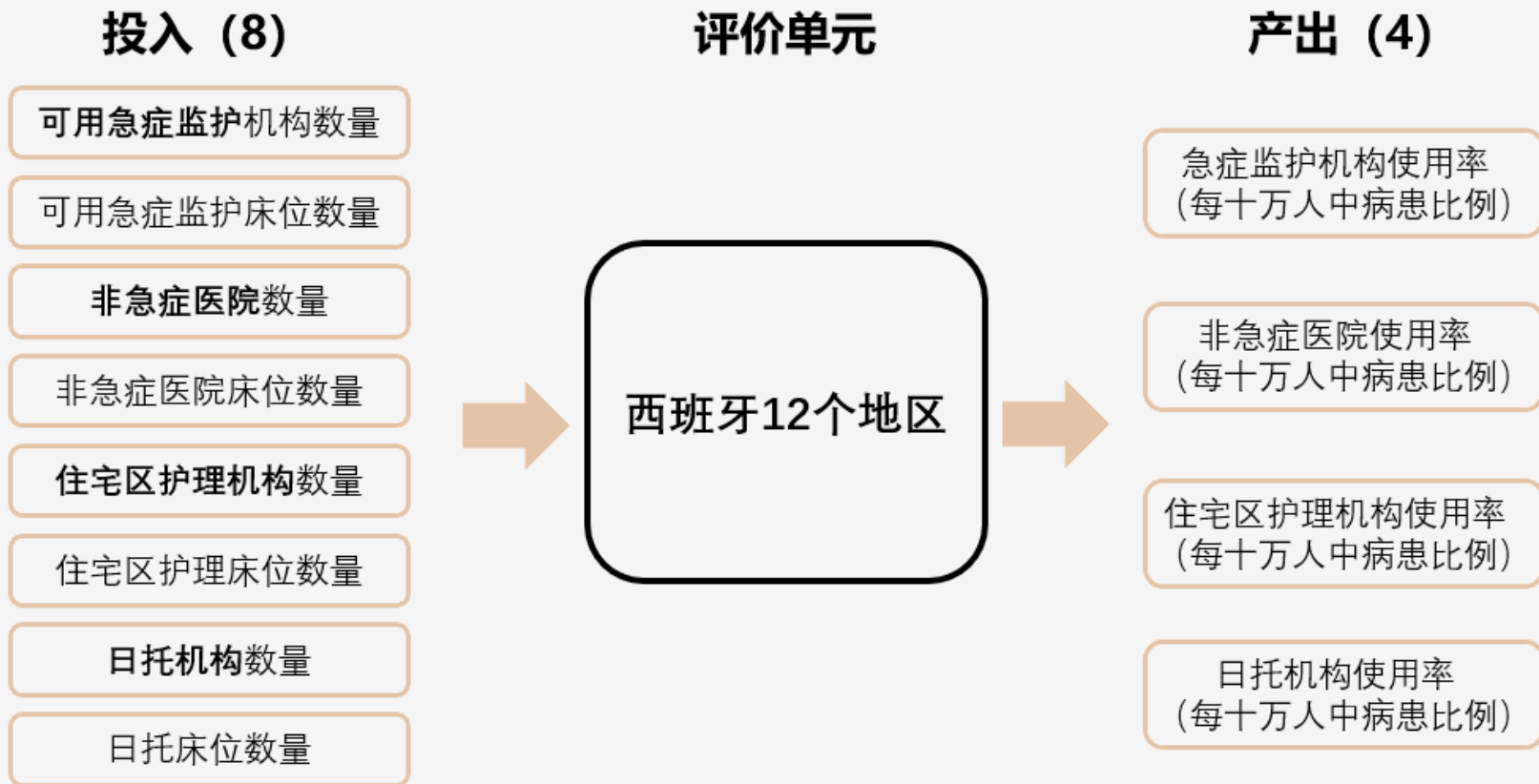
1.**指标合理性**：投入产出指标的选择、处理是否合理，符合现实含义

2.**指标随机性**：投入产出指标值可能是随机的，因为它们有些来源于人为评价/预测，或有缺失、错误值

3.**模型区分力**：投入产出指标的数量可能远大于被评价单元的数量，导致区分力降低

1 问题来源

研究目的：设计一种适用于复杂系统（以医疗系统为例）效率评价的方法，包括专家知识和DEA模型两部分，能够考虑上述不足（指标合理性、随机、区分力）。



2 研究内容

基础模型在复杂系统中应用时，存在不足，本文的解决方法：

- 指标合理性——**把专家知识转化为函数，用于对指标进行标准化处理**
- 指标存在随机性——**结合蒙特卡罗模拟，拟合数据特征**
- 指标过多，模型区分力不足——**选择低相关的指标，组成多个“场景”，多角度评价**

方法步骤

1. 构建蒙特卡罗DEA模型
2. 选择低相关的投入产出指标组成多个场景
3. 将专家知识整合入蒙特卡罗DEA模型
4. 蒙特卡罗DEA模型中的统计误差估计
5. 模型有效性验证

2 研究内容

(1) 构建蒙特卡罗DEA模型

$nsim$: 仿真次数, n : 决策单元个数, $nsim_{min}$: 用于计算统计误差的仿真次数最小值, 即至少要执行多少次仿真, ste_{min} : 统计误差最小值, 达到后停止程序。

从 i 从1到 $nsim$:

通过蒙特卡罗仿真随机确定I/O值 (专家指定分布)

如果I/O值需要基于专家规则库修正

对相应的I/O进行变换(例如, 使用线性单调递减变换)

对每个决策单元 (从1到 n):

设定相应DEA模型, 求解模型, 保存结果

如果当前仿真次数 $\geq nsim_{min}$:

计算统计误差 Ste

如果统计误差 Ste 小于 ste_{min} : 程序结束

否则: 继续执行循环

随机确定I/O值

根据专家规则对I/O
值进行变换

设计DEA模型求解,
得到一组效率值

计算统计误差

2 研究内容

(2) 选择低相关的投入产出指标组成多个场景

思路：现有投入产出指标是 8×4 ，从中选择 3×2 的子集，称为一个场景，每个场景下可以运用蒙特卡罗DEA模型计算一次效率。

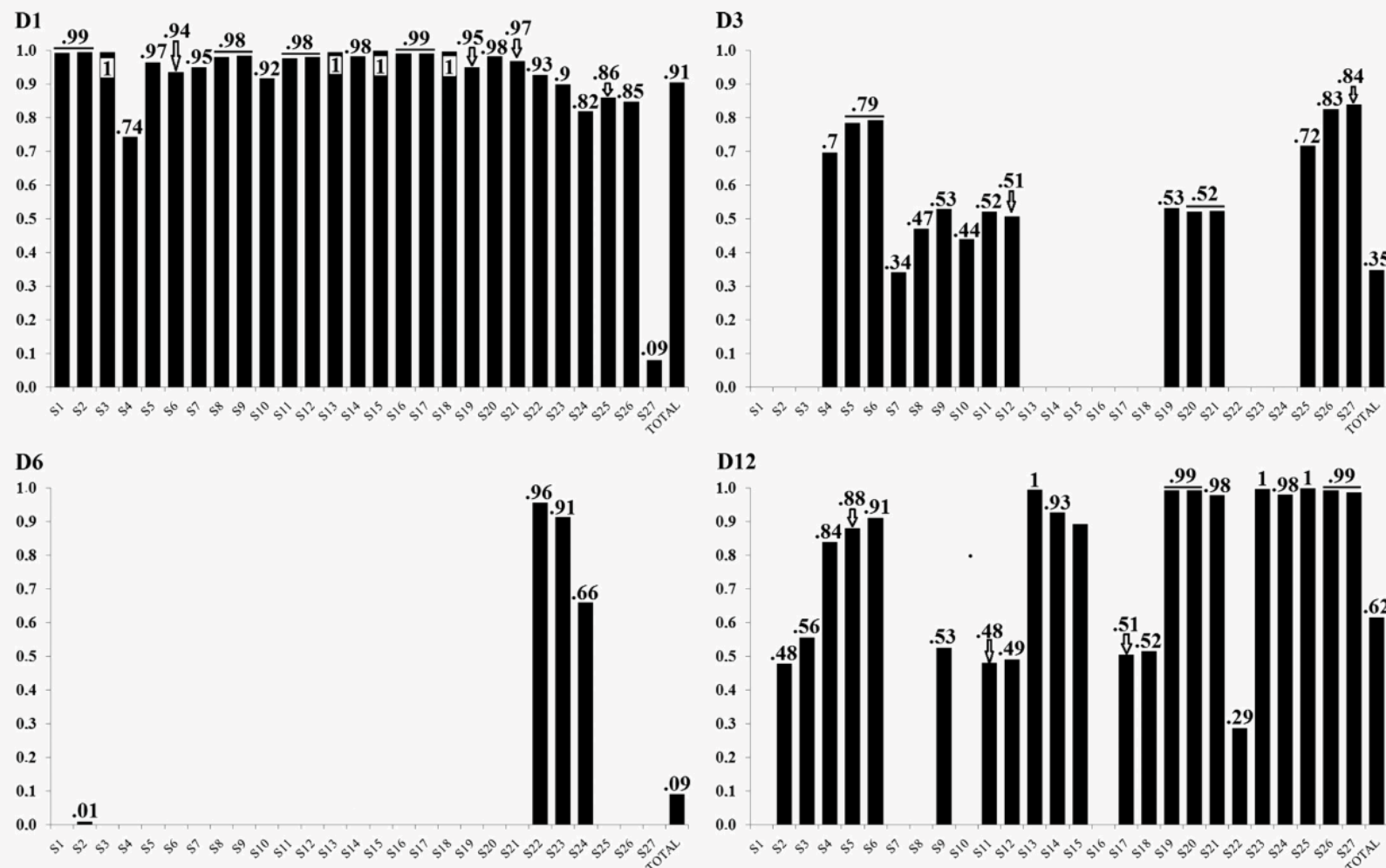
全排列组合会产生336个场景，计算量过大，只考虑指标之间相关性低的场景，共有27个。

相关性计算：

随机产生数千次 $12 \times 8 \times 4$ 的指标矩阵，任选两个指标，如果**Pearson相关系数大于0.95的仿真次数超过95%**，且**平均相关系数大于0.9**，则认为两个指标高度相关。

2 研究内容

(2) 选择低相关的投入产出指标组成多个场景



四个DMU的评价结果
横坐标：27个场景
纵坐标：效率为1的概率

2 研究内容

(3) 将专家知识整合入蒙特卡罗DEA模型

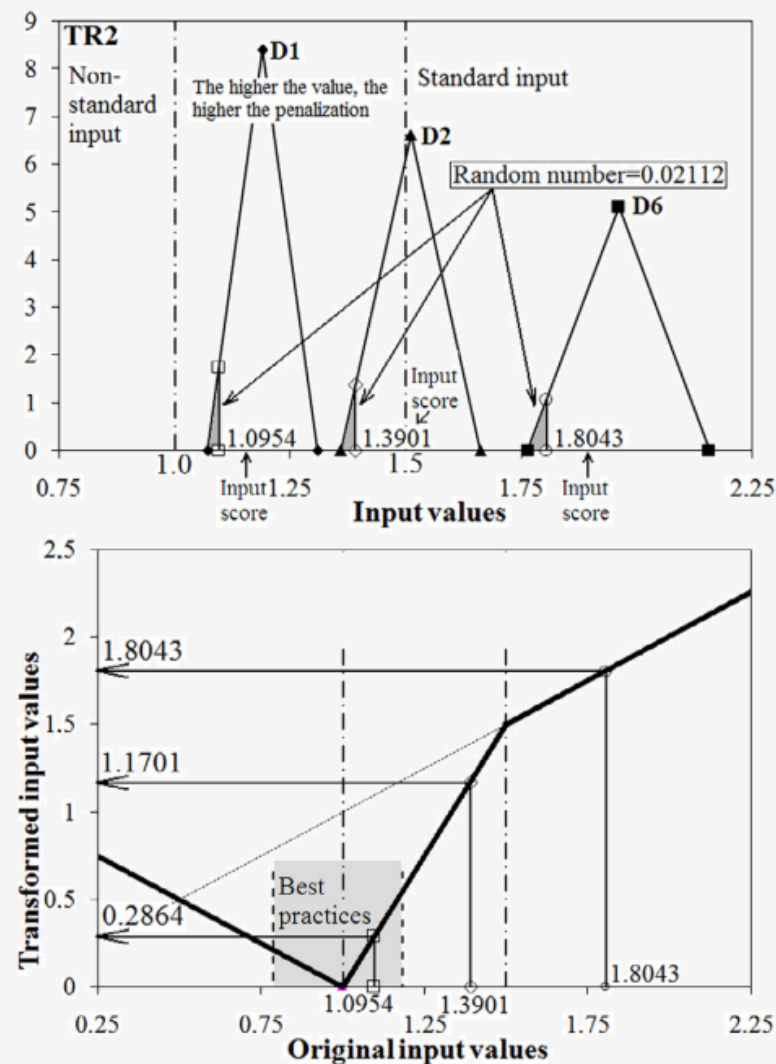
专家知识体现为：

- 为每个指标选择合适的分布形式
- 每个指标的标准化处理方式

例：投入指标“可用急症监护机构数量”拟合为三角函数。
这是一个非常规变量，因为卫生规划人员认为接近1的值是非常合适的；但作为投入指标，模型会认为它越小越好。所以需要制定规则，对原始数据进行转换。

$$x_i^{\text{tr}} = x_i^{\text{or}} \otimes f_i$$

标准化指标值=原始指标值*专家规则函数



2 研究内容

(4) 蒙特卡罗DEA模型中的统计误差估计

计算统计误差，是为了确定仿真循环的终止条件。由于没有独立同分布假设，所以不能直接估计方差作为统计误差，采用**多次重复方法**(Nakayama, 2008)来计算。

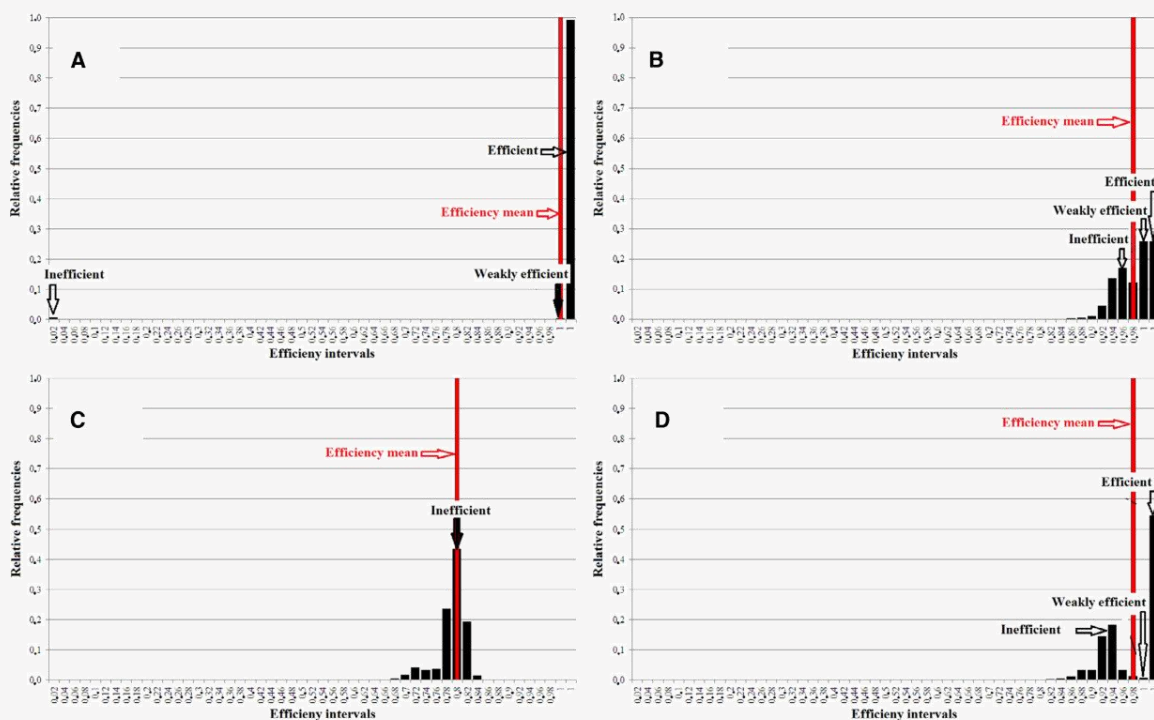
思路：不是对一个仿真函数执行多次，而是通过生成不重叠的随机数序列，产生 r 个 ($10 \leq r \leq 30$) **独立**同分布的仿真副本，每个执行 k 次。 k 需要足够大，才能估计长期运行的情况。对每个副本，可以计算出 k 次运行的平均值，共得到 r 个平均值，再计算 r 个平均值之间的方差和置信区间。

	Simulation Output			
replication 1	$Y_{1,1}$	$Y_{1,2}$	\cdots	$Y_{1,k}$
replication 2	$Y_{2,1}$	$Y_{2,2}$	\cdots	$Y_{2,k}$
\vdots	\vdots	\vdots	\ddots	\vdots
replication r	$Y_{r,1}$	$Y_{r,2}$	\cdots	$Y_{r,k}$

2 研究内容

(5) 模型有效性验证

蒙特卡罗DEA模型内部是黑箱，又已经融入了专家知识，没有其他附加信息很难验证效果。为了验证模型效果，请专家将12个决策单元分为**4组：有效、接近有效、无效和不确定的**。同时，蒙特卡罗DEA模型计算出了每个决策单元在不同场景下有效的概率，使用k均值聚类，分为4组。



四个DMU的效率分布情况

横轴：效率值，从0到1

纵轴：相对频率

红色：效率平均值

2 研究内容

(5) 模型有效性验证

	Global efficiency ^a	Experts ^b	Cluster ^c
D1	0.9060	2	1
D2	0.4068	2	2
D3	0.3487	3	2
D4	0.1612	3	3
D5	0.2920	3	3
D6	0.0902	3	3
D7	0.1488	3	3
D8	0.2597	3	3
D9	0.4451	4	4
D10	0.9888	1	1
D11	0.2832	3	2
D12	0.6163	4	4
Efficiency of the system ^f	0.3956		

专家分组

1=有效

2=接近有效

3=无效;

4=不确定

聚类分组

1=有效, 全局效率大于0.9

2=接近无效, 全局效率在0.28-0.41

3=无效, 全局效率在0.09-0.3

4=略微无效, 全局效率在0.43-0.62, 方差很大

两种分组大体一致，略有区别（相关系数0.85）

- 专家认为“接近有效”的类别，在聚类分析中实际上是低效的(D1除外)。这组决策单元的效率很低，但专家们无法准确地解释这种情况。
- 专家认为“无效”的类别，在聚类分析中也是“无效”的(除了D3和D11是接近无效)。
- 专家的“不确定”类别在聚类分析中是“略微无效”，由于其方差大，这个组也可以称为“不确定”。

所以，专家可以精准识别非常无效的单元，但中间部分的略微无效的单元对他们来说比较复杂。

3 问题解决效果、创新点

讨论：DEA模型结果是否有必要与专家分类一致？

尽管专家看法可能有误，在本文的背景中认为其是基本正确的，因此两种分组结果相似是好事。

但也不能忽略差异，因为专家无法像模型一样对所有变量进行排列组合、全局分析，因此模型可能比专家更好用。融合了专家规则的模型的可用性体现在：

- 决策时并不总是有一个专家小组，模型包含了专家知识
- 决策者可以从与他们观点一致/不同的结果中学习，可以通过变量(1/0)分析一致和不一致的原因，并以随机的方式计算潜在的改进

综上，本文证明了是**可以将专家知识（规则库）与模型结合**的，用于进行复杂系统的效率评价。并且，模型得到的结果可以进一步提高专家知识。

4 启示

1. 在指标过多时，以多个“场景”来分析的思路可以借鉴。
2. 讨论了模型结果是否有必要与专家分类一致的问题：大部分论文都相信专家结果正确，而模型可以补充更多细节。因此模型结论与专家结果一致是好的。
3. 没有论述专家选择指标拟合函数的过程，但这部分对结果会有比较大影响