

团体标准

T/CARSA 4.2—2026

遥感产品算法测评 第2部分：指标体系与指标测算方法

Algorithm test of remote sensing product—

Part2: Indicator system and indicator evaluation method

2026-03-12 发布

2026-03-15 实施

中国遥感应用协会 发布

全国团体标准信息平台

目 次

前言	II
引言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 基本要求	2
5 算法测评指标体系	2
6 算法测评指标测算	3
附录 A (资料性) 算法精度指标及测算方法	6
附录 B (资料性) 算法适用性指标及测算方法	8
附录 C (资料性) 算法效率指标及测算方法	9
附录 D (资料性) 算法鲁棒性指标及测算方法	11
附录 E (资料性) 基于层次分析法的指标权重确定方法	14
参考文献	16

前 言

本文件按照GB/T 1.1-2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件是T/CARSA 4-2026《遥感产品算法测评》的第2部分，T/CARSA4-2026已经发布了以下部分：

- 第1部分：测评流程；
- 第2部分：指标体系与指标测算方法；
- 第3部分：算法测评接口。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国遥感应用协会提出并归口。

本文件起草单位：中国科学院空天信息创新研究院、西南交通大学、浙江师范大学、鹏程实验室、西南大学、北京师范大学、北京大学、电子科技大学、武汉大学、中国科学院成都山地灾害与环境研究所、中国科学院西北生态环境资源研究院、中国林业科学院资源信息研究所。

本文件主要起草人：游冬琴、闻建光、吴小丹、林兴稳、唐勇、朴森、陈曦、孙韬、赵娜、刘照言、王新鸿、肖青、柳钦火、宫宝昌、李丽、欧阳晓莹、刘强、马明国、肖志强、范闻捷、吴骅、何涛、赵伟、车涛、田昕、魏秋芳。

引 言

随着遥感技术的发展及应用普及，遥感产品算法逐渐广泛应用于产品生产。为了提高遥感产品生产水平，保障遥感产品生产效率和产品精度，制定科学、统一的遥感产品算法测评规范，已成为遥感行业的重要任务。算法是遥感产品的核心，全面评估算法效能有助于改进算法精度和提高产品质量。

通过构建全面的评价体系和指标测算方法，规范测评方法和流程，使遥感产品测评人员在各种遥感产品算法测评时有据可依，从而提高测评的质量和效率，推动遥感产品算法的持续优化，更好地服务遥感产品生产。

T/CARSA 4-2006《遥感产品算法测评》旨在确立普遍适用于遥感产品算法测评的流程和实施步骤，拟由三个部分构成。

——第1部分：测评流程。目的在于为遥感产品算法测评提供可操作步骤与顺序。

——第2部分：指标体系与指标测算方法。目的在于为遥感产品算法构建测评指标体系及其测算方法。

——第3部分：算法测评接口。目的在于为用于测评的遥感产品算法接口提供编制规则，规范算法集成接口。

遥感产品算法的复杂性要求算法测评工作亟需建立科学、系统的评价规范。当前，针对算法评价虽存在多种指标，但缺乏统一的测评指标体系和规范化的量化流程，导致测评结果难以直接比较，同时影响了对算法功能和性能的准确全面的判断。本文件是T/CARSA 4-2006《遥感产品算法测评》的第2部分，旨在与第1部分确立的测评流程框架紧密配合，系统构建测评指标体系。本文件凝练了科学合理的测评指标体系，明确了指标的计算方法，规范了指标量化过程，为第1部分中“指标体系构建”提供具体、可操作的实施依据，共同保障测评过程的严谨性、测评结果的全面性、可比性和可重复性。同时，本部分所确立的指标体系和测算方法也可独立使用，为其他领域的算法或者产品评价工作提供参考。

遥感产品算法测评 第2部分：指标体系与指标测算方法

1 范围

本文件规定了遥感产品算法测评的指标体系构建基本要求、指标体系和指标测算方法。
本文件适用于应用于遥感产品生产算法的测评指标体系构建，其他算法测评也可参考构建。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 32453-2015 卫星对地观测数据产品分类分级规则
GB/T 36296-2018 遥感产品真实性检验导则
JJF 1059.1-2012 测量不确定度评定与表示
T/CARSA 4.1-2026 遥感产品算法测评：第1部分 测评流程

3 术语和定义

GB/T 32453-2015、GB/T 36296-2018、JJF 1059.1-2012、T/CARSA 4.1-2026界定的以及下列术语和定义适用于本文件。

3.1

算法精度 algorithm accuracy

算法生产的遥感产品的测试值与参考数据之间的一致性程度。

注：包括准确度和不确定度。

3.2

算法适用性 algorithm applicability

算法在反演不同特征（如地表类型、地形、时空尺度）目标下满足遥感产品生产需求的能力。

3.3

算法效率 algorithm efficiency

算法在特定的运行环境下，生产出特定的遥感产品所需的运行时间和计算资源。

3.4

算法鲁棒性 algorithm robustness

算法在不同运行环境和输入条件下保持产品精度和生产效率稳定的能力。

4 基本要求

遥感产品算法测评指标体系构建应符合T/CARSA 4.1-2026 第4章的规定，还应符合以下基本要求：

- a) 测评体系应包含指标和量化方法；
- b) 应涵盖一级指标，二级指标按测评需求选择；
- c) 指标测算结果应量化为统一的值域。

5 指标体系

5.1 指标体系构成

算法测评指标体系包括一级指标和二级指标两个层级，见表1。

表 1 算法测评指标层级

一级指标	二级指标
算法精度	准确度
	不确定度
算法适用性	地表类型适用性
	地形特征适用性
	空间尺度适用性
算法效率	运算时间效率
	运算空间效率
	辅助数据依赖性
算法鲁棒性	算法容错性
	算法稳定性
	时空完整性

5.2 算法精度

5.2.1 准确度

算法输出结果与相对真值间的一致程度，通常用平均误差、平均相对误差、均方根误差等指标予以表征。

5.2.2 不确定度

算法输出结果的分散性，指示结果的可信性、有效性的怀疑程度或不确定程度，通常用方差、标准不确定度、合成标准不确定度等指标予以表征。

5.3 算法适用性

5.3.1 地表类型适用性

算法在不同地表类型（如陆地地表覆盖分类、水体类别等）标准数据下生产出满足精度要求的遥感产品的能力。

5.3.2 地形特征适用性

算法在不同地形特征标准数据下生产出满足精度要求的遥感产品的能力。

5.3.3 空间尺度适用性

算法在不同空间分辨率标准数据下生产出满足精度要求的遥感产品的能力。

5.4 算法效率

5.4.1 时间效率

算法在单位时间内生产的遥感产品数据量。

5.4.2 空间效率

算法使用单位存储空间资源（内存、硬盘空间）生产的遥感产品数据量。

5.4.3 辅助数据依赖性

算法对辅助数据输入的依赖程度。

5.5 算法鲁棒性

5.5.1 算法容错性

算法在输入标准数据中设定的错误或异常数据时，仍能正常运行的能力。

5.5.2 算法稳定性

算法在输入标准数据中含有不同程度噪声或误差的数据时，输出结果保持一致性和可靠性的能力，或者不同数据量输入下，算法性能保持稳定的能力。

5.5.3 时空完整性

算法在标准数据输入下运算，反演结果在时间和空间维度上有效值的比例。

6 指标测算

6.1 算法精度指标测算

6.1.1 指标计算

将算法输出结果与参考数据进行比对，计算指标。其中：

- a) 数值型产品中评价像元值的准确度指标计算方法按GB/T 36296-2018附录A执行；
- b) 几何型产品中评价像元几何准确度指标计算方法见附录A.1；
- c) 类别型产品中评价准确度指标如误差矩阵、总体分类精度、Kappa系数，计算方法见GB/T 36296-2018附录A；漏检率和错分率,计算方法见附录A.2；
- d) 不确定度只适用于数值型产品，常用指标及计算方法按GB/T 36296-2018附录B.2执行。

6.1.2 精度测算

依据遥感产品的精度要求设定测算评定阈值，将精度指标数值定量化转为测评的分数值，测算方法见附录A.3。

6.2 算法适用性指标测算

6.2.1 指标计算

输入不同特征目标（地表类型、地形特征、空间尺度）的标准数据进行产品生产，依据输出结果满足精度要求的特征数量（地表类型数量、地形特征数量、空间尺度量级）确定适用性，指标计算方法见B.1。

6.2.2 适用性测算

将适用性指标数值定量化转为测评分数值，测算方法见附录B。

6.3 算法效率指标测算

6.3.1 指标计算

在相同标准数据输入下运行算法，统计指标。其中：

a) 时间效率

统计单位时间算法生产的遥感产品数据量，指标计算见附录C.1.1。

b) 空间效率

统计单位存储空间资源（内存、硬盘空间）算法生产的遥感产品数据量，指标计算见附录C.1.2。

c) 辅助数据依赖性

测试和统计算法输入的辅助数据种类、必要性和复杂度计算，若算法明确必须依赖的辅助数据，则直接计入统计的数量，指标计算见附录C.1.3。

6.3.2 效率测算

将效率指标数值定量化转为测评分数值，测算方法见附录C.2。

6.4 算法鲁棒性指标测算

6.4.1 指标计算

在不同标准数据输入下，依据输出结果统计指标。其中：

a) 容错性

将含有错误的标准数据输入算法，统计能正常运行数据量比例，指标计算方法见附录D.1.1。

b) 稳定性

稳定性包括但不限于抗噪声稳定性和性能稳定性。抗噪声稳定性，输入含有不同噪声的标准数据，统计满足精度要求的输出结果；性能稳定性，输入不同数据量的标准数据，统计每次输出单位数据量的产品所用时间的差异。指标计算见附录D.2.1。

c) 时空完整性

输入所有标准数据，计算输出有效结果的比例，指标计算方法见附录D.3.1。

6.4.2 鲁棒性测算

将鲁棒性指标定量化转为测评分数值。其中，容错性测算见附录D.1.2，稳定性见附录D.2.2，时空完整性见附录D.3.2。

6.5 综合测算

6.5.1 指标权重确定

依据专家经验打分并确定各指标两两之间的相对重要性，赋予一级指标和二级指标权重，权重确定方法见附录E。

6.5.2 综合测算方法

将二级指标测算值依据相应权重进行线性加和测算一级指标，再将一级指标测算值依据相应权重进行线性加和测算综合值。

附 录 A
(资料性)
算法精度指标及测算方法

A.1 几何精度指标

A.1.1 高程中误差

算法反演高程与真值偏差的平方和同名点数量 n 比值的平方根, 符号为 σ , 由式 (A.1) 计算:

$$\sigma = \sqrt{\frac{1}{n} \sum_{k=1}^n (Z_k - z_k)^2} \dots\dots\dots (A.1)$$

式中:

Z_k ——数字高程模型或数字表面模型产品第 k 个待检验点高程;

z_k ——第 k 个同名点高程相对真值;

A.1.2 高程平均误差

算法反演高程与相对真值误差的平均值, 符号为 B_d , 由式 (A.2) 计算:

$$B_d = \frac{1}{n} \sum_{i=1}^n (Z_k - z_k) \dots\dots\dots (A.2)$$

式中:

Z_k ——数字高程模型或数字表面模型产品第 k 个待检验点高程;

z_k ——第 k 个同名点高程相对真值;

n ——同名点数量。

A.1.3 平面位置中误差

数字正射影像平面位置与相对真值误差平方和与同名点数量 n 比值的平方根, 符号为 M , 由式 (A.3) 计算:

$$M = \sqrt{\frac{1}{n} \sum_{i=1}^n (\Delta x_i^2 + \Delta y_i^2)} \dots\dots\dots (A.3)$$

式中:

Δx_i ——第 i 个同名点在平面上 x 方向的误差;

Δy_i ——第 i 个同名点在平面上 y 方向的误差。

A.2 类别精度指标

A.2.1 漏检率

测定真实类别为 i 而分为其它类别的概率, 第 i 类的漏检率用 OE_i 表示, 由式 (A.4) 计算:

$$OE_i = \frac{P_{+i} - n_{ii}}{P_{+i}} \dots\dots\dots (A.4)$$

式中:

P_{+i} ——实际类别 (参考数据) 为第 i 类的样本数;

n_{ii} ——表示分类结果为第 i 类而实际类别为第 i 类的样本数目;

A.2.2 误检率

测定分类结果为第 j 类，而真实类别为其它类的概率，第 j 类的误检率用 CE_j 表示，由式 (A.5) 计算：

$$CE_j = \frac{P_{j+} - n_{jj}}{P_{j+}} \dots\dots\dots (A.5)$$

式中：

P_{j+} ——分类结果为第 j 类的样本数；

n_{jj} ——表示分类结果为第 j 类而实际类别为第 j 类的样本数目；

A.3 测算方法

A.3.1 阈值设置

依据遥感产品精度要求，设置精度阈值 $[A_{\min}, A_{\max}]$ ，按指标满足阈值范围赋予相应分数 $[S_{\min}, S_{\max}]$ ，如最低分 $S_{\min}=50$ ，最高分 $S_{\max}=100$ 。

A.3.2 测算公式

若指标值越小精度越高(如绝对偏差，均方根误差)，按公式 A.6 测算；否则，按公式 A.7 测算。

$$f_{acc} = \begin{cases} S_{\min}, & a \geq A_{\max} \\ S_{\min} + \frac{A_{\max} - a}{A_{\max} - A_{\min}} \times (S_{\max} - S_{\min}), & A_{\min} < a < A_{\max} \\ S_{\max}, & a \leq A_{\min} \end{cases} \dots\dots\dots (A.6)$$

式中：

f_{acc} ——精度指标测算值（分数值）；

a ——精度指标值；

A_{\max} ——精度指标值阈值最大值；

A_{\min} ——精度指标值阈值最小值；

S_{\max} ——最高分值（如 100）；

S_{\min} ——最低分值（如 50）。

$$f_{acc} = \begin{cases} S_{\min}, & a \leq A_{\min} \\ S_{\min} + \frac{a - A_{\min}}{A_{\max} - A_{\min}} \times (S_{\max} - S_{\min}), & A_{\min} < a < A_{\max} \\ S_{\max}, & a \geq A_{\max} \end{cases} \dots\dots\dots (A.7)$$

式中：

f_{acc} ——精度指标测算值（分数值）；

a ——精度指标值；

A_{\max} ——精度指标值阈值最大值；

A_{\min} ——精度指标值阈值最小值；

S_{\max} ——最高分值（如 100）；
 S_{\min} ——最低分值（如 50）。

附录 B (资料性) 算法适用性指标及测算方法

B.1 适用性指标

适用性指标按照公式B.1计算：

$$A = \frac{\sum_{i=1}^N a_i}{N} \dots\dots\dots (B.1)$$

式中：

A ——算法适用性指标值，如地表覆盖、地形等；

N ——测试情境总数；

a_i ——在给定的测评情境的标准输入下(i)，依据算法结果是否满足精度要求判定适用性，按照式(B.2)赋值，

$$a_i = \begin{cases} 1 & \text{算法满足精度要求} \\ 0 & \text{算法不满足精度要求} \end{cases} \dots\dots\dots (B.2)$$

B.2 测算方法

算法适用性按式B.3测算。

$$f_{\text{fea}} = A \times S_{\max} \dots\dots\dots (B.3)$$

式中：

f_{fea} ——适用性测算值（分数值），如地表覆盖、地形等；

A ——适用性指标值，按式 B.1 赋值；

S_{\max} ——最高分值（如 100）。

附录 C
(资料性)
算法效率指标及测算方法

C.1 算法效率指标

C.1.1 时间效率指标采用式C.1计算：

$$E_t = \frac{C_{output}}{T} \dots\dots\dots (C.1)$$

式中：

- E_t ——算法的时间效率值；
- C_{output} ——输出产品的总数据量；
- T ——从算法运行到输出产品所需总时间。

C.1.2 空间效率指标采用式C.2计算：

$$E_s = \frac{C_{output}}{a_1 \cdot C_{RAM} + a_2 \cdot C_{Disk}} \dots\dots\dots (C.2)$$

式中：

- E_s ——算法的空间效率值；
- C_{output} ——输出产品的总数据量；
- a_1 ——内存成本权重，可用单位容量的硬件成本作为权重；
- a_2 ——磁盘资源的成本权重，可用单位容量的硬件成本作为权重；
- C_{RAM} ——算法得到输出产品的过程中所占用的最大内存资源；
- C_{Disk} ——算法得到输出产品的过程中所占用的最大磁盘存储资源。

C.1.3 辅助数据依赖性指标采用式C.3计算：

$$E_{aux} = \sum_{i=1}^N a_i \cdot c_i \dots\dots\dots (C.3)$$

式中：

- E_{aux} ——算法的辅助数据依赖性值；
- N ——测评的辅助数据类型总数；
- a_i —— a_i 算法对第*i*类辅助数据必要性判断,可按式C.4赋值；
- c_i ——第*i*类辅助数据复杂度赋分, c_i 可根据经验确定,依据第*i*类辅助数据的获取成本或使用难度进行赋分。

$$a_i = \begin{cases} 2 & \text{算法必要辅助数据或先验信息} \\ 1 & \text{算法次要辅助数据或先验信息} \end{cases} \dots\dots\dots (C.4)$$

C.2 测算方法

C.2.1 时间效率和空间效率测算方法

C.2.1.1 阈值设定

时间效率和空间效率指标测算采用动态阈值的方法。在测算过程中，参考前一次算法测评（同参数）测算得到的指标数值作为确定本次指标阈值；当为首次测评时，将满足精度指标的指标数值最高值作为本次测评的上限阈值 E_{\max} ，最低值作为本次测评的下限阈值 E_{\min} ，随着后续算法的测评更新指标阈值范围，选择满足精度指标的指标数值最高值作为上限阈值 E_{\max} 。

C.2.1.2 效率测算

时间效率或空间效率指标按照公式C.5测算：

$$f_{eff_ss} = \begin{cases} S_{\max}, & (E \geq E_{\max}) \\ S_{\min} + \frac{E - E_{\min}}{E_{\max} - E_{\min}} \times (S_{\max} - S_{\min}), & (E_{\min} < E < E_{\max}) \\ S_{\min}, & (E \leq E_{\min}) \end{cases} \dots\dots\dots$$

(C.5)

式中：

- f_{eff_ss} ——效率指标测算分值，可为时间效率或空间效率；
- E_{\min} ——本次算法测评指标值的下限阈值；
- E_{\max} ——本次算法测评指标值的上限阈值；
- S_{\max} ——最大分值（如 100）；
- S_{\min} ——最低分值（如 0）；
- E ——效率指标值，可为时间效率或空间效率。

C.2.2 辅助数据依赖性测算方法

C.2.2.1 阈值设定

辅助数据依赖性指标的阈值设定，可依据经验设定上限阈值 E_{\max} ，也可在本次测评时，取所有参评算法该项指标分值最大值为上限阈值 E_{\max} ，无辅助数据依赖设为0。

C.2.2.2 效率测算

辅助数据依赖性指标按照公式C.6测算：

$$f_{eff_aux} = \begin{cases} S_{\max}, & (E_{aux} = 0) \\ S_{\min} + \frac{E_{\max} - E_{aux}}{E_{\max}} \times (S_{\max} - S_{\min}), & (0 < E_{aux} < E_{\max}) \\ S_{\min}, & (E_{aux} \geq E_{\max}) \end{cases} \dots\dots\dots$$

(C.6)

式中：

- f_{eff_aux} ——辅助数据依赖性指标测算分值；

- E_{\max} ——本次算法测评指标值的上限阈值；
 S_{\max} ——最大分值（如 100）；
 S_{\min} ——最低分值（如 0）；
 E_{aux} ——算法的辅助数据依赖性值。

附录 D (资料性) 算法鲁棒性指标及测算方法

D.1 容错性测算

D.1.1 指标计算

输入含错误的数量为 N 的标准数据，算法可正常执行数量为 n ，容错性指标 T_{error} 按照公式 D.1 计算：

$$T_{error} = n/N \quad \dots\dots\dots (D.1)$$

D.1.2 测算方法

容错性分值按照公式 D.2 测算：

$$f_{rob_error} = T_{error} \times S_{\max} \quad \dots\dots\dots (D.2)$$

式中：

- f_{rob_error} ——容错性指标测算分值；
 S_{\max} ——最高分值（如 100）。

D.2 稳定性测算

D.2.1 指标计算

D.2.1.1 噪声稳定性

输入含有一定等级噪声水平的噪声源数据量为 m_i 的标准输入数据，输入结果满足精度要求的数量为 n_i ，则对于第 i 类噪声的稳定性指标 T_{noise_i} 按照公式 D.3 计算：

$$T_{noise_i} = \frac{n_i}{m_i} \quad \dots\dots\dots (D.3)$$

式中：

i ——噪声类型，可按需设置，对于卫星遥感数据通常分为系统噪声、随机噪声、杂讯噪声、压缩噪声、压缩噪声、模糊噪声等 5 类。

D.2.1.2 性能稳定性

在相同的运算环境下，输入数据量为 m_i 的标准输入数据，统计使用算法生产出满足要求的遥感产品所需的时间 t_i ，则单位数据量的处理时间为 $c_i = t_i/m_i$ 。以不同数据量（依次增加）的标准输入数据 m_1, m_2, \dots, m_N 作为输入，分别统计单位数据量的处理时间 c_1, c_2, \dots, c_N ；然后，按照公式 D.4 依次计算每增加单位数据量处理时间变化量 a_i ，具体记录过程如表 D.2 所示；最后，稳定性指标 S_{eff} 按照公式 D.5 统计 a_i 的均值表示。

$$a_i = (c_{i+1} - c_i) / (m_{i+1} - m_i) \quad \dots\dots\dots (D.4)$$

表 D.2 稳定性测评过程记录

输入数据量 m_i	m_1	m_2	m_3	m_4	...	m_i	...	m_N
所需时间 t_i	t_1	t_2	t_3	t_4	...	t_i	...	t_N
$c_i = t_i/m_i$	c_1	c_2	c_3	c_4	...	c_i	...	c_N
处理速率变化量 a_i	—	a_1	a_2	a_3	...	a_i	...	a_n

$$S_{eff} = \frac{1}{N-1} \sum_{i=1}^{N-1} a_i \quad \dots\dots\dots (D.5)$$

式中:

S_{eff} ——稳定性指标值, 当 $S_{eff}=0$, 表明算法对单位数据量的处理效率不变; 当 $S_{eff}>0$, 说明随数据量增加效率降低; 当 $S_{eff}<0$, 说明随数据量增加效率提高;

a_i ——处理速率变化量;

D. 2.2 测算方法

D. 2.2.1 噪声稳定性测算方法

将各类噪声的稳定性 T_{noise_i} 进行加权求和获取总的容错性, 容错性分值按照公式D.6测算:

$$f_{rob_noise} = \sum w_i T_{noise_i} \times S_{max} \quad \dots\dots\dots (D.6)$$

式中:

f_{rob_noise} ——噪声稳定性指标测算分值;

w_i ——第 i 类噪声的权重, 可以通过文献调研给定或者专家打分给定;

T_{noise_i} ——第 i 类噪声的稳定性指标值;

S_{max} ——最高分值 (如 100)。

D. 2.2.2 性能稳定性测算方法

性能稳定性分值按照公式 D.7 测算:

$$f_{rob_eff} = \begin{cases} S_{max}, & (S_{eff} \leq 0) \\ \left(1 - \frac{S_{eff}}{\max(a_i)}\right) \times S_{max}, & (S_{eff} > 0) \end{cases} \quad \dots\dots\dots (D.7)$$

式中:

f_{rob_eff} ——性能稳定性指标测算分值;

a_i ——处理速率变化量;

$\max(a_i)$ ——统计表 D.2 中 a_i 的最大值;

S_{eff} ——稳定性指标值。

S_{max} ——最高分值 (如 100)。

D. 3 时空完整性测算

D. 3.1 指标计算

输出时空序列有效结果数量 n (结果满足产品的物理有效值范围) 占总结果数量 N 的比例 r_v , 按照公式D.8计算:

$$r_v = n/N \quad \dots\dots\dots (D.8)$$

D.3.2 测算方法

时空完整性分值按照公式D.9测算：

$$f_{rob_compl} = r_v \times S_{max} \quad \dots\dots\dots (D.9)$$

式中：

f_{rob_compl} ——时空完整性指标测算分值；

r_v ——时空完整性指标值；

S_{max} ——最高分值（如 100）。

附录 E

(资料性)

基于层次分析法的指标权重确定方法

基于层次分析法（Analytic Hierarchy Process，简称AHP）计算指标权重包括以下4个步骤：

E.1 构建层次结构

算法测评的层次结构包括目标层、准则层和措施层三个层次，可依据需要再扩展，如二级指标下再细分指标，可再措施层下分措施子层：

- a) 目标层——算法测评；
- b) 准则层——算法测评的一级指标；
- c) 措施层——算法测评的二级指标。

E.2 构建判断矩阵

E.2.1 判断矩阵

判断矩阵是反映下层影响因子对于上层影响因子的相对重要性。若将A作为上层因素，A支配的下层元素记为 $A_1, A_2, A_3, \dots, A_n$ 。例如：当A为算法测评， A_n 为一级指标；当A为一级指标， A_n 为该一级指标下的二级指标，依此类推。元素 A_i 与 A_j 的相对重要性为 a_{ij} ，则二级指标相对重要性判断矩阵 $B = (a_{ij})_{n \times n}$ 。

B 为正互反矩阵，满足 $a_{ij} > 0; a_{ij} = \frac{1}{a_{ji}}$ ($i, j=1, 2, \dots, n$)。 a_{ij} 通过E.2.2专家打分确定。

E.2.2 专家打分

判断矩阵通过专家打分，量化各指标相对于其他指标的重要性。邀请多名多领域专家根据“1-9标度法”，按表E.1所示，对同一层次的指标进行两两比较，确定 a_{ij} 填写判断矩阵。

表 E.1 “1-9 标度法”含义

标度	含义
1	A_i 和 A_j 同等重要
3	A_i 比 A_j 重要一点
5	A_i 比 A_j 更重要
7	A_i 比 A_j 重要得多
9	A_i 比 A_j 极端重要
2, 4, 6, 8	用于更细致的描述
倒数	若 A_i 与 A_j 的重要性之比为 a_{ij} ，则 A_j 与 A_i 的重要性之比为 $\frac{1}{a_{ij}}$

E.3 权重计算

判断矩阵B的权重向量，可按如下步骤计算：

- a) 按式E.1计算判断矩阵B中每个元素的积 M_i 。

$$M_i = \prod_{j=1}^n a_{ij} \quad (i = 1, 2, \dots, n) \quad \dots\dots\dots (E.1)$$

- b) 按式E.2计算每行 M_i 的n次方根 $\bar{\omega}_i$ 。

$$\bar{\omega}_i = \sqrt[n]{M_i} \quad \dots\dots\dots (E.2)$$

- c) 将向量 $\bar{\omega} = (\bar{\omega}_1, \bar{\omega}_2, \dots, \bar{\omega}_n)^T$ 进行归一化处理，按式E.3计算出权重系数 ω_i 。

$$\omega_i = \frac{\bar{\omega}_i}{\sum_{i=1}^n \bar{\omega}_i} \quad \dots\dots\dots (E.3)$$

E.4 一致性检查

在求得判断矩阵 $B = (a_{ij})_{n \times n}$ 中各指标的权重后，进行一致性检验，即指对B确定不一致的允许范围。具体步骤如下：

- a) 按照公式E.4计算一致性指数CI，CI值越大，一致性越差。

$$CI = \frac{\lambda_{\max} - n}{n - 1} \quad \dots\dots\dots (E.4)$$

式中：

n ——判断矩阵阶数；

λ_{\max} ——判断矩阵最大特征根，按式E.5计算。

$$\lambda_{\max} = \frac{1}{n} \sum_{i=1}^n \frac{(B\omega)_i}{\omega_i} \quad \dots\dots\dots (E.5)$$

式中：

n ——判断矩阵阶数；

B——判断矩阵；

ω_i ——权重系数；

ω ——权重系数构成的权重向量；

- b) 根据阶数 $n (\leq 10)$ 从表E.2查找对应的随机一致性指标RI (Random Index)，处理超过 10 维的判断矩阵时，可借助专业软件支持或者模拟计算等。

表 E.2 随机一致性指标 RI 的标准值

n	1	2	3	4	5	6	7	8	9	10
RI	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

- c) 按照公式E.6计算一致性比例CR (Consistency Ratio)来确定判断矩阵一致性。若 $CR < 0.1$ ，则判断矩阵具有满意的一致性；若 $CR \geq 0.1$ ，则不满足，则需要进一步调整判断矩阵，直到 $CR < 0.1$ 为止。

$$CR = \frac{CI}{RI} \quad \dots\dots\dots (E.6)$$

参 考 文 献

- [1] GB/T 3358.1-2009 统计学词汇及符号 第1部分：一般统计术语与用于概率的术语
[2] GB/T 3358.2-2009 统计学词汇及符号 第2部分：应用统计
-

全国团体标准信息平台