

## 基于近红外光谱的设施栽培水果黄瓜磷元素亏缺初期快速诊断

石吉勇<sup>1</sup>, 邹小波<sup>1\*</sup>, 赵杰文<sup>1</sup>, 毛罕平<sup>2</sup>, 王开亮<sup>1</sup>, 陈正伟<sup>1</sup>, 黄晓玮<sup>1</sup>

1. 江苏大学食品与生物工程学院, 江苏 镇江 212013
2. 江苏大学现代农业装备与技术重点实验室, 江苏 镇江 212013

**摘要** 磷元素(P)亏缺初期, 水果黄瓜植株根部叶片出现小斑点, 其症状的外观特征与健康植株根部叶片老化初期类似, 难以用肉眼或者计算机图像处理技术识别。本文根据近红外光谱能够反映叶片组织中有有机物组分的差异, 运用近红外光谱技术对水果黄瓜植株磷元素亏缺进行了快速诊断研究。精确控制营养液中磷元素含量, 通过设施栽培方式培养缺磷植株和对照样本。近红外光谱仪采集了90片叶子的原始光谱(60片作为训练集, 30片作为预测集), 经光谱预处理和窗口宽度优化后均匀划分为27个子区间, 分别提取每个子区间的10个主成分数据作为BP人工神经网络(BP-ANN)的输入变量, 以叶片缺素情况作为输出变量, 建立3层BP-ANN诊断模型。当主成分因子数为3时, 第7个子区间对应的模型效果最佳, 模型对缺磷叶片和正常叶片的预测准确率均达到100%。研究表明: 近红外光谱技术结合BP-ANN快速诊断水果黄瓜磷元素亏缺是可行的。

**关键词** 缺素诊断; 近红外光谱技术; 磷元素; 水果黄瓜; BP人工神经网络

中图分类号: O657.3 文献标识码: A DOI: 10.3964/j.issn.1000-0593(2011)12-3264-05

### 引言

设施栽培作物以其早熟、高产、增效快等优点受到农民普遍欢迎, 是最有活力的农业新兴产业<sup>[1]</sup>。由于其生长周期短、产出量大、需肥量多, 要求肥力供给具有高消耗、高补充的特点, 生产中设施栽培作物很容易出现氮、磷、钾比例失调和缺素症状, 其中磷是植物体内核酸、核蛋白、磷脂、ATP和辅酶等重要化合物的组分, 能加强光合作用和促进氮素代谢<sup>[2,3]</sup>。磷元素亏缺容易导致作物营养不良, 严重影响了作物的产量和质量以及经济效益的进一步提高, 因此有必要对设施栽培作物的磷元素亏缺进行快速、精确诊断<sup>[4]</sup>。

磷元素的常规检测法为化学方法测定植物组织或者其提取物中磷元素的总含量, 该方法不仅费时费力, 而且会损害采样植株<sup>[5]</sup>。同时, 磷元素在植株中分有机磷和无机磷两种形态, 其分布在细胞及植物组织内有明显的区域化现象, 虽然化学方法能够准确测定磷元素的含量, 但是测量过程没有考虑磷元素的形态和分布, 因此难以准确反映植株的磷元素营养状况<sup>[6]</sup>。为了克服上述不足, 近年来无损检测技术在

作物营养状况检测中得到了广泛运用, 如毛罕平等采用机器视觉技术对番茄缺素进行了诊断, 取得了较好的识别效果<sup>[7-9]</sup>。Marian Wiwart采用彩色图像分析技术对三种豆类的大量元素亏缺进行了研究, 详细分析了缺素叶片图像的色差差异<sup>[10]</sup>。

本研究在控制磷元素供给不足条件下, 以温室无土栽培方式培育水果黄瓜缺磷植株。通过观察缺磷植株生长过程发现, 缺磷初期植株根部附近老叶首先出现小斑点, 随着缺磷条件下生长时间推移小斑点逐渐扩大并伴随组织坏死。外观上缺磷初期叶片病症同老叶老化初期的特征类似, 难以用肉眼或者计算机图像技术进行辨别。前人的研究表明作物在缺素状态下, 叶片组织的化学组分会发生变化, 考虑近红外光谱能够反映叶片组织中有有机物组分的差异, 因此本研究用近红外光谱技术对缺磷初期的黄瓜植株进行诊断<sup>[11]</sup>。作物叶片的组分非常复杂, 对应的近红外光谱异常复杂, 必须结合相应的化学计量学方法进行解析<sup>[12]</sup>。误差反向传输人工神经网络(BP-ANN)是一种有监督的学习算法, 以样本的输出与期望输出的偏差极小为目标进行学习, 利用梯度最速下降法, 使权值误差沿误差函数的负梯度方向改变, 具有分布

收稿日期: 2011-02-23, 修订日期: 2011-06-18

基金项目: 国家(863计划)项目(2008AA10Z208), 国家自然科学基金项目(60901079), 江苏省六大大才高峰和青蓝工程项目, 国家博士后基金项目, 优秀博士论文基金项目和江苏省研究生科研创新计划项目(CX10B\_277Z)资助

作者简介: 石吉勇, 1984年生, 江苏大学食品与生物工程学院博士研究生 e-mail: stoneboy\_2007@vip.sohu.com

\*通讯联系人 e-mail: zou\_xiaobo@ujs.edu.cn

并行处理、非线性映射、自适应学习和稳健容错等特性,被广泛用于光谱检测中的非线性建模<sup>[13]</sup>。在实际运用过程中通常采用主成分分析算法将光谱数据压缩为几个变量,将其作为神经网络的输入变量<sup>[14,15]</sup>。

为了提高计算效率,本研究将光谱均匀划分为若干个子区间,分别提取每个子区间的若干主成分数据作为 BP-ANN 的输入变量,以叶片缺素情况作为输出变量,建立 3 层 BP-ANN 磷元素亏缺诊断模型,尝试快速对磷元素亏缺初期的水果黄瓜植株进行诊断,为及时追肥、挽救生产提供科学指导。

## 1 实验部分

### 1.1 样本培育

水果黄瓜品种为碧玉 3 号(上海富农种业有限公司提供),以珍珠岩(镇江培蕾有机肥有限公司提供)为基质,在温室大棚内通过无土栽培方式培养水果黄瓜植株。营养液选用日本的山崎配方<sup>[16]</sup>,水果黄瓜生长初期用正常营养液浇灌,在水果黄瓜植株长出 4~5 片叶子后采用缺磷营养液进行浇灌,缺磷营养液是在正常营养液的基础上完全扣除磷元素。为了保证试验结果的可靠性,分别用缺磷营养液培育了 40 株缺磷水果黄瓜样本,用正常营养液培育了 40 株参照水果黄瓜样本。

### 1.2 近红外光谱数据采集

采摘缺磷植株的缺素初期叶片以及正常植株的老化初期叶片采集近红外光谱数据。数据采集设备为 Antaris II 型傅里叶变换近红外光谱仪(Thermo Fisher, 美国),采用 In-GaAs 检测器,波数范围 10 000~4 000  $\text{cm}^{-1}$ ;扫描次数为 32 次;波数间隔为 1.928  $\text{cm}^{-1}$ ;分辨率为 16  $\text{cm}^{-1}$ 。数据采集过程中,温度保持在 25  $^{\circ}\text{C}$  左右,湿度保持基本不变。每个样品采集图 2 圆圈内斑点区域的近红外光谱数据,每个斑点区域采集三次并取三次数据的平均值作为原始光谱。试验共采集了 90 条近红外光谱(缺磷样本 50 条光谱,正常样本 40 条光谱),60 条作为训练集,剩余 30 条作为预测集。

### 1.3 数据降维和波段选择

样本的近红外光谱包含 3 112 个波数点,信息量非常大。如果将光谱数据直接作为输入变量进行建模,不仅会因为变量太多导致计算效率低,而且还会引入噪声降低模型的预测精度。为了有效提取近红外光谱的特征信息,采用主成分分析(PCA)现实光谱数据降维。国内外学者比较全光谱偏最小二乘法(PLS)和不同区间偏最小二乘法(iPLS, BiPLS, Si-PLS)建立的近红外光谱模型,发现 iPLS 模型的预测效果往往优于全光谱 PLS 模型,说明选用信噪比高的波段比全光谱对应的模型效果更好<sup>[17]</sup>。因此,本研究将全光谱划分为若干个子区间,交互验证了主成分因子数和特征子区间对模型识别率的影响,并得到最优光谱诊断模型。光谱数据处理流程如图 1 所示。

### 1.4 神经网络建模

神经网络是一种重要的模式识别方法,其中多层误差反向传播神经网络方法(back-propagation, BP)应用尤广,

具有强大的非线性建模能力,特别适合解决复杂的映射问题<sup>[18,19]</sup>。水果黄瓜叶片是组分复杂的生物组织,叶片的理化性质与其可见/近红外反射光谱之间是一种复杂的映射关系,本研究将特征波段降维后的数据作为输入变量,采用 BP-ANN 算法建立叶片磷元素亏缺的诊断模型。

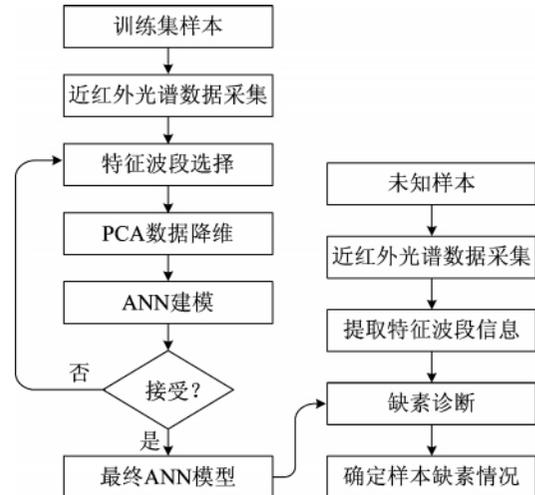
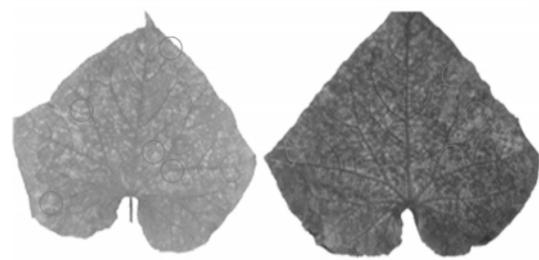


Fig 1 Flowchart of data processing

## 2 结果与讨论

### 2.1 缺磷样本及正常样本比较

水果黄瓜在磷元素亏缺初期,植株根部老叶出现小斑点,如图 2(a)中小圆圈区域所示。正常植株老叶在开始老化时,其叶面也会出现小斑点,如图 2(b)中的小圆圈区域所示。从图 2 中可以看出,缺磷初期叶片和正常老化初期对应的特征斑点无论在颜色还是在形状上都十分类似,难以用肉眼或者计算机图像技术进行及时、快速诊断。



(a) 缺磷初期黄瓜叶片 (b) 老化初期黄瓜叶片

Fig 2 Phosphorus-deficient leaf and health leaf

虽然磷元素亏缺初期叶子的症状同正常植株老叶老化症状在外观上类似,但是这些症状的诱导因素不一样,必然导致缺磷对应斑点区域和老叶老化斑点区域内叶片的有机物组成、含量和比例不一致。前人的大量研究表明,近红外光谱能够很好的反映有机物组成、含量等方面的差异,非常利于对磷元素亏缺初期的黄瓜叶片进行快速诊断。缺磷样本和正常样本斑点区域的近红外光谱如图 3 所示,其中实线为缺磷叶片对应的光谱,虚线为正常叶片对应的光谱。从图 3 中可以看出,两者的近红外光谱存在较大的差异,这些差异有利

于模式识别算法诊断样本是否缺乏磷元素。

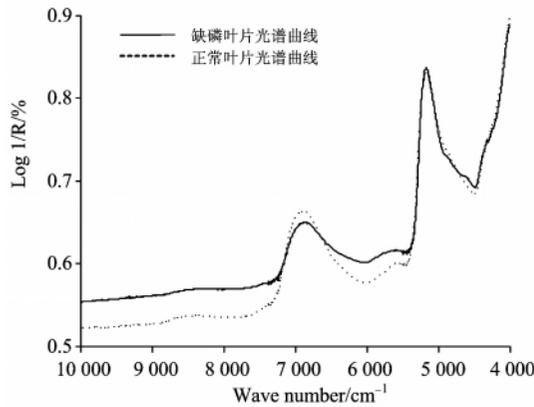


Fig 3 Near infrared spectra of phosphorus-deficient leaf and health leaf

## 2.2 参数优化及 BP-ANN 建模

比较图 3 中两类样本的光谱曲线差异可以看出, 两条光谱在不同波数下的差别程度不一致, 因此将全光谱进行 ANN 建模难以得到最优建模结果。为了提高模型效果并降低计算量, 本工作将全光谱均匀分为若干子区间, 分别考察每个子区间对应模型的识别效果, 并选择预测效果最好的模型作为最优模型。在本试验中, BP-ANN 模型的输入层单元数是主成分因子数; 它的输出层单元数为 1, 即叶片是否缺磷; 选择 S 型函数作为模型的作用函数; 训练过程中的学习因子和动量因子都置为 0.1, 模型中间层单元数通过训练的结果来优化。BP-ANN 建模计算过程中发现子区间窗口宽度对模型的识别效果影响较大, 需要对其进行优化, 结果如图 4 所示。图 4 中  $x$  轴为光谱子区宽度, 其取值范围为 [88 194] (对应的子区间总数为 [16 35]);  $y$  轴表示当前子区间总数对应的最佳 ANN 模型识别率。从图 4 可以看出, 子区间窗口宽度为 115 和 148 个波数点时, 模型识别率均达到 100%。为了使模型尽量简单, 选择 115 作为最佳子区间窗口宽度。

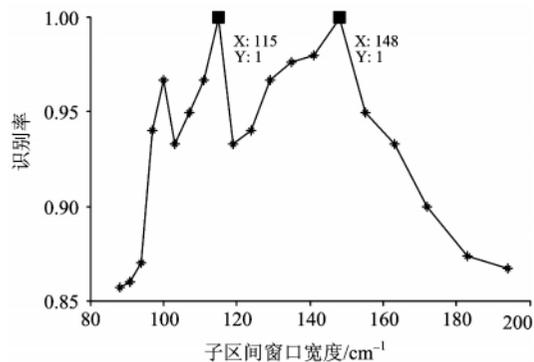


Fig 4 Optimal result of total interval width

区间宽度为 115 时(此时全光谱被均匀划分为 27 个子区间), 各个子区间对应的 ANN 模型训练集识别率如图 5 所示。图 5 中  $x$  轴表示子区间数,  $y$  轴表示 ANN 模型识别率, 柱状图底部数值表示 ANN 模型的最佳主成分因子数, 图中两条光谱曲线分别为缺磷叶片近红外光谱和正常叶片近红外

光谱。比较图 5 中两条光谱曲线与训练集识别率的关系, 可以发现两条光谱曲线差异大的子区间对应的模型识别率较高, 如前 12 个子区间对应的识别率基本大于 0.9, 表明这些区间对应的近红外光谱包含了大量磷元素亏缺信息。从图 5 可以看出, 第 6, 7, 8 子区间的训练集识别率均为 1, 预测集识别率分别为 1, 1 和 0.96, 最佳主成分因子数分别为 7, 3 和 6。由于模型识别率越高, 代表建模效果越好; 主成分因子数越小, 代表建模所用变量数越少, 建立的模型越简单, 本文选择第 7 子区间作为最优子区间, 其对应的 BP-ANN 模型为最优诊断模型。

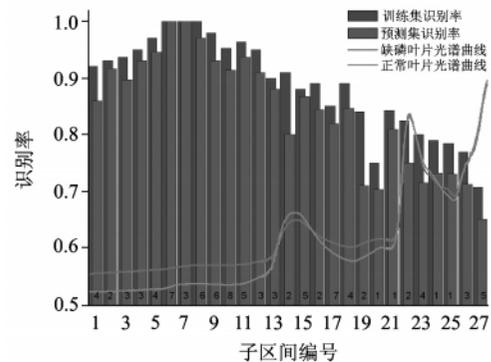


Fig 5 Recognition rate of ANN model under different principal component factors and intervals

## 2.3 模型比较

为了客观评价 ANN 诊断模型的建模效果, 本研究采用 KNN 和 ANN 方法对全光谱建立了缺素诊断模型, 结果如表 1 所示。ANN 全光谱模型预测集识别率为 95%, 优于 KNN 全光谱模型, 表明 ANN 模式识别方法具有较强的识别能力。区间 ANN 模型同全光谱 KNN 模型相比, 训练集和预测集识别率均达到 100%, 建模所需的波数点大大降低, 同时最优主成分因子数也小于全光谱 ANN 模型, 表明区间 BP-ANN 能够有效的筛选特征波数点, 提高模型识别率, 并降低模型复杂度。

Table 1 Discriminating results of different models

模型	波数点	Pcs	训练集识别率/%	预测集识别率/%
全光谱 KNN	3 112	7	90	87
全光谱 BP-ANN	3 112	8	95	93
区间 BP-ANN	115	3	100	100

注: Pcs: 最佳主成分因子数

## 3 结束语

水果黄瓜磷元素亏缺初期, 叶片表现出的缺素症状同正常叶片老化初期症状类似, 难以用肉眼或者计算机图像处理技术进行识别。为了对磷元素亏缺初期的植株进行快速诊断, 本研究采集了叶片病症区域的近红外光谱, 结合 BP-ANN 算法建立了磷元素亏缺初期的光谱诊断模型。建模时将全光谱划分为 27 个子区间, 第 7 子区间的前 3 个主成分

数据对应的 BP-ANN 模型最优, 模型对训练集的正确识别率为 100%, 预测集的识别率为 100%。研究表明近红外光谱技术结合 BP-ANN 算法能够对磷元素亏缺初期的植株进行快速诊断。缺磷植株培育实验表明, 水果黄瓜植株磷元素亏

缺后期, 叶片组织大面积坏死, 严重影响植株生长, 因此在磷元素亏缺初级阶段的快速诊断植株营养状况, 对及时指导追肥和挽救生产具有重要意义。

## References

- [ 1 ] CAO Wei-xing, ZHU Yan, TIAN Yong-chao, et al(曹卫星, 朱艳, 田永超, 等). *Scientia Agricultura Sinica*(中国农业科学), 2006, (2): 281.
- [ 2 ] Milton N M, Eiswerth B A, Ager C M. *Remote Sensing of Environment*, 1991, 36(2): 121.
- [ 3 ] YAO Q, YANG K, PAN G, et al. *Agricultural Sciences in China*, 2007, 6(5): 559.
- [ 4 ] ZHANG Li-xia, PENG Jian-ming, MA Jie(张丽霞, 彭建明, 马洁). *Chinese Agricultural Science Bulletin*(中国农学通报), 2010, (8): 157.
- [ 5 ] LU Jian-ping, PENG Jian, SHI Jian-rong, et al(陆建平, 彭剑, 石建荣, 等). *Chinese Journal of Analytical Chemistry*(分析化学), 2008, 36(2): 238.
- [ 6 ] Johnston M E, Gikaara D M, Edwards D G. *Scientia Horticulturae*, 2006, 110(3): 298.
- [ 7 ] MAO Han-ping, WU Xue-mei, LI Ping-ping(毛罕平, 吴雪梅, 李萍萍). *Transactions of the Chinese Society of Agricultural Engineering*(农业工程学报), 2005, 21(8): 106.
- [ 8 ] MAO Han-ping, XU Gui-li, LI Ping-ping(毛罕平, 徐贵力, 李萍萍). *Transactions of the Chinese Society of Agricultural Machinery*(农业机械学报), 2003, 34(2): 73.
- [ 9 ] MAO Han-ping, XU Gui-li, LI Ping-ping(毛罕平, 徐贵力, 李萍萍). *Transactions of the Chinese Society of Agricultural Engineering*(农业工程学报), 2003, 19(2): 133.
- [ 10 ] Wiwart M, Fordonski G, Zuk-Golaszewska K, et al. *Computers and Electronics in Agriculture*, 2009, 65(1): 125.
- [ 11 ] Petisco C, García-Criado B, de Aldana B R V, et al. *Analytical and Bioanalytical Chemistry*, 2005, 382(2): 458.
- [ 12 ] Li X, He Y, Fang H. *Journal of Food Engineering*, 2007, 81(2): 357.
- [ 13 ] CHANG Wei-wei, GUO Lei, FU Zhao-yang, et al(常威威, 郭雷, 付朝阳, 等). *Journal of Infrared and Millimeter Waves*(红外与毫米波学报), 2010, 29(3): 205.
- [ 14 ] Dou Y, Sun Y, Ren Y, et al. *Analytica Chimica Acta*, 2005, 528(1): 55.
- [ 15 ] Zou X B, Zhao J W, Povey M J W, et al. *Analytica Chimica Acta*, 2010, 667(1-2): 14.
- [ 16 ] AN Na, XU Hui, SUN Zhou-ping, et al(安娜, 须晖, 孙周平, 等). *Journal of Shenyang Agricultural University*(沈阳农业大学学报), 2006, 37(3): 495.
- [ 17 ] Zou X B, Zhao J W, Li Y X. *Vibrational Spectroscopy*, 2007, 44(2): 220.
- [ 18 ] SHAN Yang, ZHU Xiang-rong, XU Qing-song, et al(单杨, 朱向荣, 许青松, 等). *Journal of Infrared and Millimeter Waves*(红外与毫米波学报), 2010, 29(2): 128.
- [ 19 ] WU Yan, WAN Wei(武妍, 万伟). *Journal of Infrared and Millimeter Waves*(红外与毫米波学报), 2007, 26(1): 65.

## Rapid Diagnostics of Early Phosphorus Deficiency in Mini-Cucumber Plants under Protected Cultivation by Near Infrared Spectroscopy

SHI Ji-yong<sup>1</sup>, ZOU Xiao-bo<sup>1\*</sup>, ZHAO Jie-wen<sup>1</sup>, MAO Han-ping<sup>2</sup>, WANG Kai-liang<sup>1</sup>, CHEN Zheng-wei<sup>1</sup>, HUANG Xiao-wei<sup>1</sup>

1. School of Food and Biological Engineering, Jiangsu University, Zhenjiang 212013, China

2. Key Laboratory of Modern Agricultural Equipment and Technology, Jiangsu University, Zhenjiang 212013, China

**Abstract** The morphological symptom of phosphorus deficiency at early stage is similar to the appearance of leaf aging process in preliminary phase, so that visual diagnostics of phosphorus deficiency in mini-cucumber plants at early stage is practically impossible. Near infrared reflectance spectra contain information about differences in compositions of leaf tissues between phosphorus-deficient plants and healthy plants. In the present paper, near infrared reflectance spectroscopy was used to provide diagnostic information on phosphorus deficiency of mini-cucumber plants grown under non-soil conditions. Near infrared spectra was collected from 90 leaves of mini-cucumber plants. Raw cucumber spectra was preprocessed by SNV and divided into 27 intervals. The top 10 principal components (PCs) were extracted as the input of BP-ANN classifiers by principal component analysis

(PCA) while the values of nutrient deficient were used as the output variables of BP-ANN and three layers BP-ANN discrimination model was built. The best experiment results were based on the top 3 principal components of No. 7 interval when the spectra was divided into 27 intervals and identification rates of the ANN model are 100% in both training set and the prediction set. The overall results show that NIR spectroscopy combined with BP-ANN can be efficiently utilized for rapid and early diagnostics of phosphorus deficiency in mini-cucumber plants.

**Keywords** Diagnostics of deficiency; Near infrared spectroscopy; Phosphorus; Mini-cucumber plants; BP-ANN

(Received Feb. 23, 2011; accepted Jun. 18, 2011)

\* Corresponding author

## 《光谱学与光谱分析》对来稿英文摘要的要求

来稿英文摘要不符合下列要求者，本刊要求作者重写，这可能要推迟论文发表的时间。

1. 请用符合语法的英文，要求言简意明、确切地论述文章的主要内容，突出创新之处。
2. 应拥有与论文同等量的主要信息，包括四个要素，即研究目的、方法、结果、结论。其中后两个要素最重要。有时一个句子即可包含前两个要素，例如“用某种改进的 ICP-AES 测量了鱼池水样的痕量铅”。但有些情况下，英文摘要可包括研究工作的主要对象和范围，以及具有情报价值的其他重要信息。在结果部分最好有定量数据，如检测限、相对标准偏差等；结论部分最好指出方法或结果的优点和意义。
3. 句型力求简单，尽量采用被动式，通常应有 2000 个印刷字符，300 个英文单词为宜，不能太短；也不要太长。用 A4 复印纸单面隔行打印。
4. 摘要不应有引言中出现的内容，换言之，摘要中必须写进的内容应尽量避免在引言中出现。摘要也不要对论文内容作解释和评论，不得简单重复题名中已有的信息；不用非公知公用的符号和术语；不用引文，除非该论文证实或否定了他人已发表的论文。缩略语、略称、代号，除相邻专业的读者也能清楚地理解外，在首次出现时必须加以说明，例如用括号写出全称。