

# 基于模糊聚类的海河流域农业节水分区研究

丁志宏<sup>1</sup>，郭兵托<sup>2</sup>，杨朝瀚<sup>3</sup>

(1.天津市中水科技咨询有限责任公司,天津 300170; 2.黄河勘测规划设计有限公司,河南 郑州 450003;  
3.水利部海河水利委员会,天津 300170)

**摘要:**结合流域机构对海河流域农业节水灌溉管理工作的现实需求,在综合考虑海河流域现状农业节水工作各项影响因素和制约条件的基础上,构建了海河流域农业节水分区评价指标体系,运用模糊聚类方法将海河流域划分为6个农业节水一级分区、19个农业节水二级分区,讨论了分区合理性,并结合最严格水资源管理制度的实施提出了下一步的工作方向。

**关键词:**海河流域; 农业节水; 节水分区; 指标体系; 模糊聚类; 最严格水资源管理制度

**中图分类号:** TV213 **文献标识码:** A

海河流域水资源匮乏,多年平均水资源总量 370 亿  $m^3$ ,人均水资源量 270 $m^3$ ,仅为全国人均水资源量的 12.8%,是全国七大江河流域中的人均水资源量最少的流域。而与此同时,海河流域又是我国经济社会发达地区,也是粮食主产区之一,社会经济与生产生活用水量巨大,农业灌溉用水量占总用水量的比重较大。据统计,海河流域 2012 年农业用水量为 252 亿  $m^3$ ,占总用水量的 67.77%,其中,农田灌溉用水量 221.85 亿  $m^3$ ,占农业用水量的 88.03%。

随着各类各级别经济社会发展规划的陆续部署和推进实施,海河流域的工业和城市生活用水量将会有较明显的增加,“工、农、生”争水的问题将日趋严重。因此,在今后一个时期内,大力发展农业节水灌溉将是确保海河流域粮食安全和农业可持续发展的重要途径,是落实“节水优先、空间均衡、系统治理、两手发力”的新时期治水思路进而实现以水资源的可持续发展支撑社会经济可持续发展的重要战略措施和实践抓手。

但在农业节水灌溉工作开展的实践过程中,海河流域各区域在地形地貌、耕作制度、作物种类、作物需水量、水资源条件、农业生产条件以及社会经济发展水平等方面具有明显的空间差异性,为了贯彻落实最严格水资源管理制度,提高流域管理机构对农业节水灌溉工作指导的针对性,亟需根据海河流域各地的自然地理条件、社会发展状况、经济承受能力和水资源条件等制约因素来划分得出不同类型的节水灌溉分区体系,进而实现分区管理。

农业节水分区可以利用的方法较多,主要有经验法、指标法、重叠法、聚类法等<sup>[1]</sup>,而其中又以模糊聚类法的应用最为广泛<sup>[2-7]</sup>。模糊聚类法有着较为严格的理论基础和计算方法,能够解释因素之间的内在本质联系和差别,消除了定性分析的主观性和任意性,适用性较强。当资料获取较为充分时,分析结果准确、可靠,能反映客观实际情况<sup>[8]</sup>。

综上所述,本文将在综合考虑海河流域现状农业节水工作各项影响因素和制约条件的基础上,构建分区评价指标体系,运用模糊聚类模型方法,开展海河流域农业节水分区研究工作,为流域层面上的农业灌溉规划与管理提供技术参考和基础支撑。本研究开展的流域层面上的农业节水分区划分工作在全国七大江河流域中尚属首次。

基金项目:国家自然科学基金项目(51309202);天津大学水利工程仿真与安全国家重点实验室开放基金项目(HESS-1717);水利部公益性行业科研专项经费项目(201101016)。

作者简介:丁志宏(1979—),男,博士,高级工程师,主要从事水文学及水资源问题的研究与咨询工作

## 1 模糊聚类方法

模糊聚类方法是一种多元数据统计分析方法,它是基于待分类对象所具有的模糊性特征,通过建立模糊相似矩阵,运用模糊等价关系进行聚类分析。

模糊聚类方法的一般步骤包括数据标准化、标定(即构建模糊相似矩阵)、聚类、确定最佳模糊截集 $\lambda$ 值等4个步骤。

### 1.1 数据标准化

#### (1) 数据矩阵

设论域 $U = \{x_1, x_2, \dots, x_n\}$ 为被分类对象,每个对象又有 $m$ 个指标表示其性状,即

$$x_i = \{x_{i1}, x_{i2}, \dots, x_{im}\}, \quad (i = 1, 2, \dots, n) \quad (1)$$

于是,得到原始数据矩阵为:

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix} \quad (2)$$

式中, $x_{nm}$ 表示第 $n$ 个分类对象的第 $m$ 个指标的原始数据。

#### (2) 数据归一化

由于 $m$ 个指标的量纲和数量级大都是不同的,若直接利用原始数据进行计算,可能会导致某一项指标在评价过程中的影响过大以及量纲单位的改变造成分类结果的改变。为了消除各个指标量纲的差别和数量级不同的影响,需要进行数据的归一化处理。数据归一化,就是将数据压缩到区间 $[0,1]$ 上。本文采用标准差法对原始指标数据进行标准化处理。

### 1.2 标定

在数据标准化之后,需要推求模糊相似矩阵。本文采用海明距离法来确定相似系数,建立模糊相似矩阵,确定 $x_i$ 与 $x_j$ 的相似程度。

海明距离:

$$r_{ij} = 1 - c \sum_{k=1}^m |x'_{ik} - x'_{jk}| \quad (3)$$

式中, $r_{ij}$  ( $i = 1, 2, \dots, n; j = 1, 2, \dots, n$ )为模糊相似关系矩阵 $R$ 的第 $i$ 行第 $j$ 列元素, $c$ 是一个特定的正数,用以保证 $0 \leq r_{ij} \leq 1$ , $x'_{ik}$ 、 $x'_{jk}$ 分别为数据标准化以后的矩阵中的第 $i$ 行第 $k$ 列元素。

### 1.3 聚类

本文采用基于模糊等价矩阵的传递闭包法进行聚类分析。

通过标定所得的模糊相似矩阵  $R$  通常只能满足自反性和对称性, 不满足传递性, 因而还不能用以进行分类。需要从  $R$  出发, 求一个包含  $R$  的最小传递矩阵——传递闭包  $t(R)$ 。一般采用平方法求传递闭包。平方法的原理为:

$$R^2 = R \circ R \quad (4)$$

$$R^4 = R^2 \circ R^2 \quad (5)$$

$$\dots\dots \quad (6)$$

$$R^{2m} = R^{2m-1} \circ R^{2m-1} \quad (7)$$

直到第一次  $R^k = R^k \circ R^k$  出现时, 则  $R^k$  就是传递闭包  $t(R)$ 。再让  $\lambda$  由大变小, 就可以完成聚类。

#### 1.4 确定最佳阈值

根据实践经验和专家意见确定适当的  $\lambda$  值, 选择待分类对象的一个具体分类以得到符合流域农业灌溉发展实际现状的分区结果。

## 2 实际应用

### 2.1 评价单元

选取海河流域综合规划中确定的 80 个水资源三级区套地市单元作为本文研究工作的基本评价单元。其中, “子牙河山区大同”、“子牙河山区朔州”和“徒骇马颊河安阳”这 3 个分区的面积仅分别为  $10\text{km}^2$ 、 $90\text{km}^2$  和  $68\text{km}^2$ , 加之其相关的自然地理、社会经济、农业灌溉等统计数据缺乏, 从实际管理工作考虑, 本次分区评价暂不考虑这 3 个单元。故本文计算使用的基本评价单元共计 77 个。

### 2.2 评价指标体系

评价指标采用两级分类指标体系构成, 一级分区指标体系主要体现的是地形地貌、灌溉水源和水资源现状开发利用程度, 具体包括地貌类型和灌溉水源等 2 项定性指标以及水资源开发利用率和地下水开发利用程度等 2 项定量指标; 二级分区指标体系主要体现的是气候特征、社会经济发展水平、灌溉用水需求状况、农业灌溉发展水平等, 具体包括干旱指数、人均 GDP、小麦灌溉定额、玉米灌溉定额、耕地灌溉率、粮田占有有效灌溉面积比例、节水灌溉比、高效节灌比、现状灌溉水利用系数、总用水量中的灌溉用水占比等 10 项定量指标。

具体指标的意义和作用, 阐述如下:

#### (1) 地形地貌

地貌类型是一个区域地表起伏程度的宏观形态, 本文所述及的海河流域地貌类型以山区(含山间盆地)和平原区来分类, 用于节水区划分类及其命名。

#### (2) 灌溉水源

灌溉水源是指灌溉水的来源, 通常分为当地水源和外调水源, 当地水源主要是地表水和地下水, 外调水源在海河流域农业灌溉中主要是引黄水。本文所述及的灌溉水源是特指引黄水, 用于节水区划分类及其命名。

#### (3) 水资源开发利用程度

水资源开发利用程度是反映一个区域已开发利用的水资源量占水资源总量的情况, 是与

社会经济供用水量、节水水平、用水总量控制指标、用水效率指标等因素密切相关的代表性指标,是一个区域现状水资源紧缺状况的集中反映,一定程度上决定了农业节水技术的发展趋势和实施方法。

考虑到海河流域地表水水质状况和以地下水为主的供水水源结构,本文采用水资源开发利用率和地下水开发利用这 2 项指标来反映水资源开发利用程度。

#### (4) 气候特征

气候对于农业生产影响极大。一般来说,水分状况是决定农业生产条件的主要因素。蒸发量和降水量是水分状况的两个主要表征指标,二者的比值可以反映一个地区水资源的天然收支情况,间接反映该地区的干湿程度,称之为干旱指数。干旱指数相似的地区,其潜在的农业水资源供需情况大致上是相近的。因此,本文选择年蒸发量与年降水量之比——干旱指数来表征气候特征。

#### (5) 社会经济发展水平

一个区域整体的社会经济发展水平是影响区域农业节水灌溉发展水平和需求的重要因素。本文采用人均 GDP 来表征社会经济发展水平。

#### (6) 灌溉定额因素

灌溉定额是一个区域在不同降水条件下的各类农作物单位面积上的灌溉需水量,可以反映一个区域基本灌溉需求能力状况。本文采用小麦和玉米这两类海河流域的主要粮食作物的 50% 频率下的灌溉需水量来表征灌溉定额这一节水分区划分影响因素。

#### (7) 灌溉发展水平因素

耕地灌溉率是有效灌溉面积占耕地面积的比例,其高低可反映当地农业生产的水利化程度、可供灌溉用水量的多少以及水资源利用的难易程度,是一项综合性指标。种植结构是影响农业需水量和节水灌溉发展的内生变量,可以用粮田灌溉面积占有效灌溉面积的比例来进行综合性表征。节水灌溉比,即节水灌溉面积占有效灌溉面积的比例,是灌溉发展水平的基准表征指标。以微灌、喷灌、管灌等为代表的高效节水灌溉方式是农业节水灌溉发展的主要方向,高效节灌比,即高效节水灌溉面积占有效灌溉面积的比例,是农业节水灌溉发展水平高低的主要表征指标。现状灌溉水利用系数是反映现状农业水资源利用有效程度的综合性指标。用水结构是反映农业水资源丰富程度的定量指标,可以用总用水量中的灌溉用水量占比来表征。为此,本文选取耕地灌溉率、有效灌溉面积中的粮田占比、节水灌溉比、高效节灌比、现状灌溉水利用系数、总用水量中的灌溉用水量占比等 6 项指标来作为农业灌溉现状发展水平方面的表征指标。

综上所述,本文所构建的农业节水分区研究评价指标体系见图 1 所示。

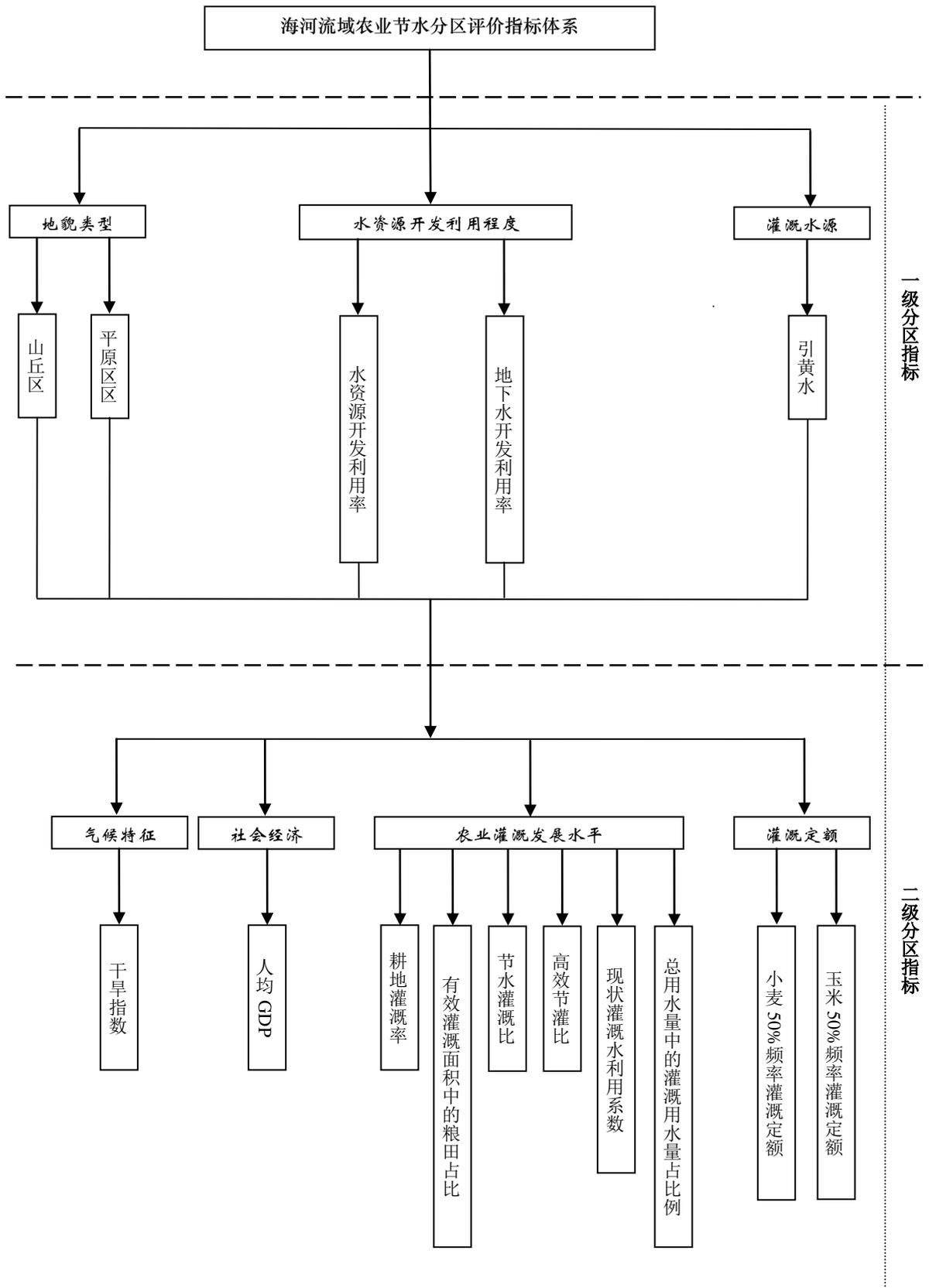


图1 海河流域农业节水分区研究的评价指标体系

### 2.3 分区划分结果

选取 2012 年为评价现状水平年, 77 个基本评价单元的 12 项定量评价指标值见表 1, 其中的水资源开发利用率为 2005~2012 年各评价单元的多年平均供水总量与其水资源总量的比值, 地下水开发利用率为 2005~2012 年各评价单元的多年平均地下水供水量与其地下水资源量的比值。

由于我国北方地区水资源短缺, 依据地表水资源合理开发利用应维持在 60%~70% 范围之内<sup>[9]</sup>以及国际公认的水资源开发利用极限合理阈值为 40% 的研究结论, 结合海河流域现状水资源开发利用实际情况, 笔者对各评价单元水资源开发利用程度指标的分类提出如下标准: (1) 水资源高强度开发利用区: 水资源开发利用率大于等于 0.6, 或者地下水开发利用率大于等于 0.6; (2) 水资源中强度开发利用区: 水资源开发利用率大于等于 0.4、小于 0.6, 或者地下水开发利用率大于等于 0.4、小于 0.6; (3) 水资源低强度开发利用区: 水资源开发利用率小于 0.4, 或者地下水开发利用率小于 0.4。

表 1 海河流域农业节水分区定量评价指标

| 水资源三级区 | 地级行政区 | 水资源开发利用率 | 地下水开发利用率 | 干旱指数 | 人均 GDP (元) | 小麦灌溉定额 (m <sup>3</sup> /亩) | 玉米灌溉定额 (m <sup>3</sup> /亩) | 耕地灌溉比 | 有效灌溉面积中的粮田占比 | 节水灌溉比 | 高效节水灌溉比 | 现状灌溉水利系数 | 总用水量中的灌溉用水占比 |
|--------|-------|----------|----------|------|------------|----------------------------|----------------------------|-------|--------------|-------|---------|----------|--------------|
| 滦河山区   | 唐山    | 0.55     | 0.51     | 1.45 | 8761       | 180                        | 90                         | 0.82  | 0.9          | 0.79  | 0.72    | 0.61     | 0.45         |
|        | 秦皇岛   | 0.17     | 0.14     | 1.37 | 14624      | 180                        | 90                         | 0.77  | 0.91         | 0.45  | 0.27    | 0.54     | 0.48         |
|        | 张家口   | 0.07     | 0.41     | 2.18 | 9707       | 180                        | 90                         | 0.2   | 0.64         | 0.67  | 0.59    | 0.41     | 0.92         |
| ∴      | ∴     | ∴        | ∴        | ∴    | ∴          | ∴                          | ∴                          | ∴     | ∴            | ∴     | ∴       | ∴        | ∴            |
| 徒骇马颊河  | 聊城    | 1.94     | 0.92     | 1.78 | 23540      | 180                        | 70                         | 0.92  | 0.80         | 0.36  | 0.31    | 0.54     | 0.79         |
|        | 滨州    | 3.39     | 0.36     | 1.85 | 29978      | 180                        | 70                         | 0.66  | 0.80         | 0.24  | 0.10    | 0.54     | 0.69         |
|        | 濮阳    | 2.63     | 2.42     | 1.81 | 32192      | 180                        | 60                         | 0.83  | 0.68         | 0.37  | 0.10    | 0.50     | 0.50         |

根据表 1, 运用模糊聚类分析法划分得到海河流域农业节水分区体系, 见表 2 和图 2~。

由表 2 可知:

(1) 海河流域可划分为 6 个农业节水一级分区, 分别为山前平原区、中部及东部滨海平原区、南部引黄平原区、北部燕山区、西北部太行山区、西部太行山区;

(2) 在 77 个分类单元中, 属于高开发利用区的单元有 36 个, 属于中开发利用区的单元有 9 个, 属于低开发利用区的单元有 32 个;

(3) 在 6 个农业节水一级分区中, 就水资源开发利用程度而言, 中部及东部滨海平原区、南部引黄平原区的各评价单元均为高; 山前平原区的各评价单元仅大清河淀西平原北京为中, 其余均为高; 北部燕山区的各评价单元普遍为中和低; 西北部太行山区均为中; 西部太行山区仅漳卫河山区焦作为高, 其余均为低;

(4) 海河流域可划分为 19 个农业节水二级分区。

表2 海河流域农业节水分区体系

| 一级分区       | 二级分区     | 所含分区单元      |       | 水资源开发利用程度 | λ值    |
|------------|----------|-------------|-------|-----------|-------|
|            |          | 水资源三级区      | 地级行政区 |           |       |
| 山前平原区      | 冀东平原区    | 滦河平原及冀东沿海诸河 | 唐山    | 高         | 0.904 |
|            |          |             | 秦皇岛   |           |       |
|            | 北京平原区    | 北四河下游平原     | 北京    | 中         |       |
|            |          |             | 北京    |           |       |
|            | 天津平原区    | 北四河下游平原     | 天津    | 高         |       |
|            |          |             | 天津    |           |       |
|            | 冀中南平原区   | 北四河下游平原     | 唐山    |           |       |
|            |          |             | 廊坊    |           |       |
|            |          | 大清河淀西平原     | 石家庄   |           |       |
|            |          |             | 保定    |           |       |
|            |          | 大清河淀东平原     | 保定    |           |       |
|            |          |             | 沧州    |           |       |
|            |          |             | 廊坊    |           |       |
|            |          | 子牙河平原       | 衡水    |           |       |
| 石家庄        |          |             |       |           |       |
| 邯郸         |          |             |       |           |       |
| 邢台         |          |             |       |           |       |
| 沧州         |          |             |       |           |       |
| 漳卫河平原      | 衡水       |             |       |           |       |
| 徒骇马颊河      | 邯郸       |             |       |           |       |
| 中部及东部滨海平原区 | 邯郸平原区    | 黑龙港及运东平原    | 邯郸    | 0.952     |       |
|            | 邢台衡水平原区  | 黑龙港及运东平原    | 邢台    |           |       |
|            | 衡水       |             |       |           |       |
| 沧州平原区      | 黑龙港及运东平原 | 沧州          |       |           |       |
| 南部引黄平原区    | 漳卫河平原区   | 漳卫河平原       | 安阳    | 0.838     |       |
|            |          |             | 鹤壁    |           |       |
|            |          |             | 新乡    |           |       |
|            |          |             | 焦作    |           |       |
|            |          |             | 濮阳    |           |       |
|            | 徒骇马颊河区   | 徒骇马颊河       | 济南    |           |       |
|            |          |             | 东营    |           |       |
|            |          |             | 德州    |           |       |
|            |          |             | 聊城    |           |       |
|            |          |             | 滨州    |           |       |
| 濮阳         |          |             |       |           |       |
| 北部燕山区      | 北部山区     | 滦河山区        | 张家口   | 低         | 0.832 |
|            |          |             | 锡林郭勒  |           |       |
|            | 北京山区     | 北三河山区       | 北京    | 中         |       |
|            | 天津山区     | 北三河山区       | 天津    |           |       |
|            | 中东部山区    | 滦河山区        | 唐山    |           |       |
| 秦皇岛        | 低        |             |       |           |       |

|               |               |               |     |   |       |
|---------------|---------------|---------------|-----|---|-------|
|               |               |               | 承德  |   |       |
|               |               |               | 朝阳  |   |       |
|               |               |               | 葫芦岛 |   |       |
|               |               | 北三河山区         | 唐山  |   |       |
|               |               |               | 张家口 |   |       |
|               |               |               | 承德  |   |       |
| 西北部太行山区       | 大同朔州山区        | 永定河册田水库以上     | 大同  | 中 | 0.921 |
|               |               |               | 朔州  |   |       |
|               | 永定河册田水库至三家店区间 | 大同            |     |   |       |
|               | 永定河册田水库以上     | 乌兰察布          |     |   |       |
| 永定河册田水库至三家店区间 |               |               |     |   |       |
| 张家口山区         | 永定河册田水库至三家店区间 | 张家口           |     |   |       |
| 西部太行山区        | 北京山区          | 永定河册田水库至三家店区间 | 北京  | 低 | 0.914 |
|               |               | 大清河山区         | 北京  |   |       |
|               | 邯郸山区          | 子牙河山区         | 邯郸  |   |       |
|               |               | 漳卫河山区         |     |   |       |
|               | 一般山区          | 永定河册田水库以上     | 忻州  |   |       |
|               |               | 大清河山区         | 石家庄 |   |       |
|               |               |               | 保定  |   |       |
|               |               |               | 张家口 |   |       |
|               |               |               | 大同  |   |       |
|               |               |               | 忻州  |   |       |
|               |               | 子牙河山区         | 石家庄 |   |       |
|               |               |               | 邢台  |   |       |
|               |               |               | 太原  |   |       |
|               |               |               | 阳泉  |   |       |
|               |               |               | 晋中  |   |       |
|               |               | 漳卫河山区         | 忻州  |   |       |
|               |               |               | 长治  |   |       |
| 晋城            |               |               |     |   |       |
| 晋中            |               |               |     |   |       |
| 安阳            |               |               |     |   |       |
| 鹤壁            |               |               |     |   |       |
| 新乡            |               |               |     |   |       |
|               | 焦作            | 高             |     |   |       |



图2 海河流域农业节水分区一级分区划分

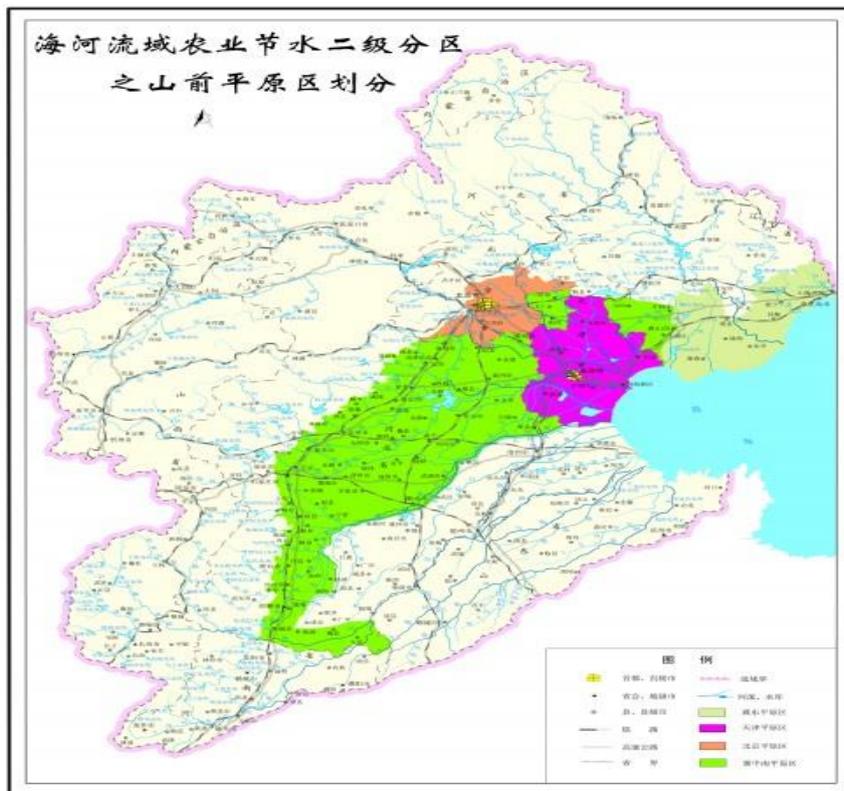


图3 海河流域农业节水二级分区之山前平原区划分

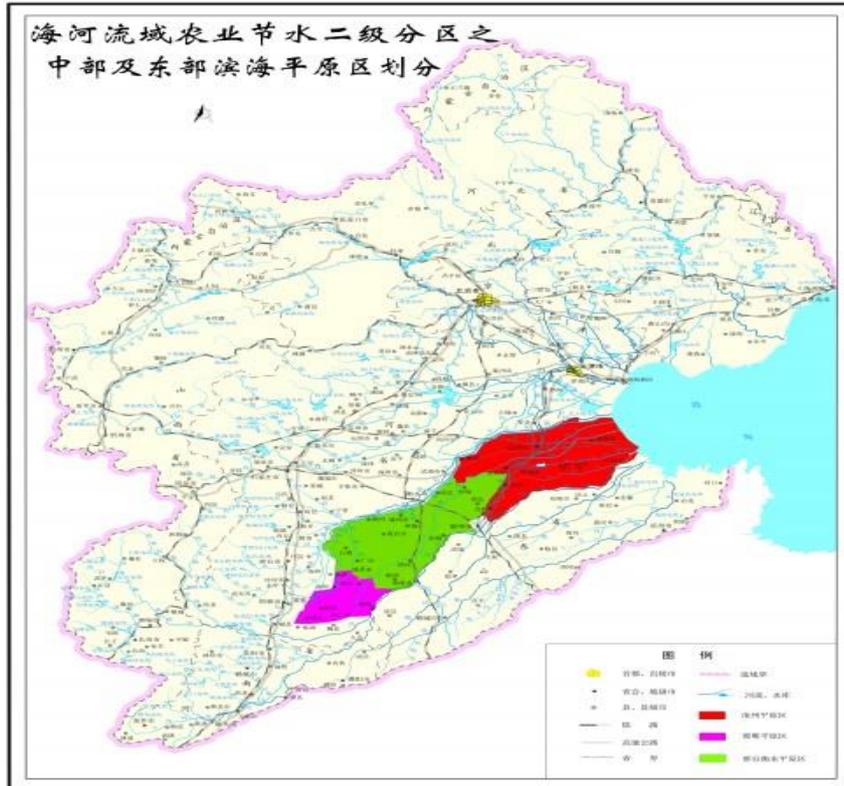


图4 海河流域农业节水二级分区之中部及东部滨海平原区划分



图5 海河流域农业节水二级分区之南部引黄平原区划分



图6 海河流域农业节水二级分区之北部燕山区划分



图7 海河流域农业节水二级分区之西北部太行山区划分

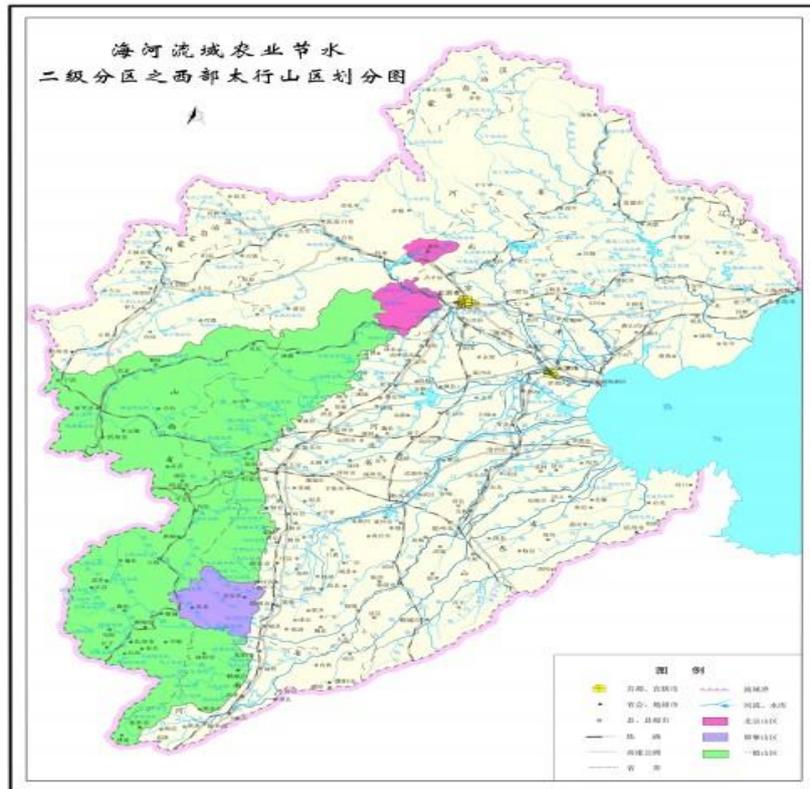


图8 海河流域农业节水二级分区之西部太行山区划分

## 2.4 分区合理性分析

### (1) 评价单元选取合理性分析

采用水资源三级区套地市分区作为基本评价单元,使得单元的空间尺度适应流域管理机构的管理需求,便于与海河流域水资源综合规划等有关规划成果相协调,便于农业节水日常管理工作的开展,便于流域机构按照最严格水资源管理制度的确定的“三条红线”、“四项制度”来开展农业节水灌溉的水资源管理工作。

### (2) 分区成果合理性分析

所划分的6个一级节水分区,既体现了与地形地貌特征有关的气候条件、作物熟制、种植结构等农业节水分区影响因素分布的空间差异性,又反映了现状海河流域水资源开发利用程度这一水资源管理主导影响因素的空间差异性,符合流域实际情况;所划分的19个二级节水分区与海河流域现状大中型灌区空间位置基本一致,与相关省(直辖市)的既有农业节水分区基本一致,框架是相互协调和统一的,既体现了区域农业节水灌溉发展现状,又体现了社会经济和行政区划对农业灌溉发展的影响,符合流域实际情况。

## 3 结论

本文构建了海河流域农业节水分区评价指标体系,运用模糊聚类方法将海河流域划分为6个农业节水一级分区、19个二级分区,为海河流域农业节水灌溉的分区管理工作提供了科学参考。

在下一步的工作中,可以继续结合最严格水资源管理制度中的不同水平年的“总量控制”和“用水效率”指标在水资源三级区套地市单元的分解结果,提出不同规划水平年各单

元适宜的灌溉水量控制指标以及不同节水灌溉模式下的适宜灌溉面积以便于流域管理机构更好地实施农业节水灌溉方面的水行政管理工作。

#### 参考文献:

- [1] 任春平, 杜敏, 李涛. 杨陵区节水灌溉分区规划[J].水土保持研究, 2002, 9(2): 15-18.
- [2] 吴景社, 康绍忠, 王景雷, 等. 基于主成分分析和模糊聚类方法的全国节水灌溉分区研究[J].农业工程学报, 2004, 20(4): 64-68.
- [3] 马立辉, 赵玲, 张会芹, 等. 模糊—动态聚类法在河北省农业节水区划中的应用[J].南水北调与水利科技, 2006, 4(1): 42-44.
- [4] 王红霞, 卢文喜. 河北省节水灌溉区划中模糊聚类分析与应用[J].工程勘察, 2006, 12: 31-35.
- [5] 王晓愚, 李占斌, 迟道才, 等. 新疆农业节水分区指标体系与方法[J].西安理工大学学报, 2008, 24(1): 89-93.
- [6] 何英, 郭玉川, 姜卉芳. 基于模态—动态聚类法的新疆节水区划探讨[J].节水灌溉, 2009, 2: 5-7, 10.
- [7] 褚琳琳. 基于因子分析与聚类分析的江苏省节水农业分区研究[J].灌溉排水学报, 2014, 33(3): 137-140.
- [8] 周开乐, 杨善林, 丁帅, 等. 聚类有效性研究综述[J].系统工程理论与实践, 2014, 34(9): 2417-2431.
- [9] 钱正英. 中国可持续发展水资源战略研究综合报告及各专题报告[M].北京: 水利水电出版社, 2001.