西藏农业科技
TIBET JOURNAL OF AGRICULTURAL SCIENCES

基于决策树与 WheatGrow 作物模型的青稞生长预测初探

吕奥博[1],董凯宁[1*],龚　澄[1],罗黎鸣[2],张思源[2],关卫星[2]

(1. 四川大学信息管理技术系,四川 成都　610064;2. 西藏自治区农牧科学院农业研究所,西藏 拉萨　850000)

摘　要:本研究致力于探究温度、土壤、水分、阳光等多因子对青稞生长的影响,并最终建立起用于指导实践及产量预测的高原青稞生长-预测模型。本研究利用西藏农牧研究院实验数据对影响青稞生长的因子进行程度划分,根据现有成熟禾本作物 WheatGrow 生长模型,带入青稞生长过程中不同程度的各影响因子拟定数据,得到传统生长模型的各因子对青稞生长的预测结果。并根据机器学习的决策树理论对传统生长模型中所涉及的各变量进行信息增益分析,按照各因子对作物生长的影响权重大小构建用于预测青稞生长及指导牧民生产种植活动的决策树模型。本研究建立起的有关影响青稞生长因子的青稞生长-预测模型,不仅实现了指导耕种实践及青稞产量预测的目的,同时还可依靠最新的青稞生长数据进行优化,进而更精准的分析青稞生长影响因子。

关键词:青稞;决策树;WheatGrow;西藏高海拔

中图分类号:S512.1　　**文献标识码**:A

Preliminary Study on Highland Barley Growth Prediction Based on Decision Tree and WheatGrow Crop Model

LV Ao-bo[1],DONG Kai-ning[1*],GONG Cheng[1], LUO Li-ming[2],ZHANG Si-yuan[2],GUAN Wei-xing[2]

(1. Sichuan University Information management system Department,Sichuan Chengdu 610064, China; 2. Institute of Agriculture, Academy of Agricultural and Animal Husbandry Sciences, Tibet Lhasa 850000, China)

Abstract:This study is devoted to exploring the effects of temperature, soil, water, sunlight and other factors on the growth of highland barley. Finally, a growth-prediction model of highland barley is established to guide practice and yield prediction. In this study, the factors affecting barley growth were classified according to the experimental data of Tibet Institute of Agriculture and Animal Husbandry. According to the existing WheatGrow growth model of mature gramineous crops, the data of various factors affecting barley growth in different degrees were drawn up, and the predictive results of the factors of traditional growth model on barley growth were obtained. According to the decision tree theory of machine learning, the information gain of the variables involved in the traditional growth model is analyzed, and the decision tree model for predicting the barley growth and guiding the production and planting activities of herdsmen is constructed according to the weight of each factor on crop growth. In this study, a barley growth-prediction model was established, which can not only guide cultivation practice and barley yield prediction, but also optimize barley growth data, and then analyze barley growth factors more accurately.

Key words:Highland barley;Decision tree;Wheat group model;High altitude of Tibet

西藏自然环境脆弱,农业生产条件恶劣,气象条件复杂。青稞的生长过程受到气候、土壤水分和营养元供给条件等多种因素的制约、胁迫[1]。而青稞是青藏高原上藏民族主要的粮食来源,其产量对西藏地区的社会经济效益有着极大的影响。青稞的绿色增产,通过促进农业资源与生产技术的有效组合,提高农业生产效益;保证了当地青稞的需要,满足了当地农牧民对青稞的需求;带动当地农牧民就业,增加群众集体经济收入为提高青稞的生产种植效率,应首先了解青稞生长过程中各影响因子发挥的作用,从而使有限的生产资源投入到最高效的生产活动当中。因此,研究及分析不同环境因子对青稞生长过程的影响十分必要。但是鉴于西藏农业化程度较低,可采集的青稞生长及环境因子数据有限,无法

收稿日期:2019-08-22

基金项目:2018 年科技部重点研发专项:高海拔边境地区农牧业关键技术研究与示范-子课题:高海拔边境地区青稞绿色增产增效关键技术研究与集成示范资助

作者简介:吕奥博(1999-),男,学士,主要研究方向为信息管理,E-mail:15210191737@163.com;*为通讯作者:董凯宁(1976-),男,博士,讲师,研究方向:信息管理。

使用传统的数据分析技术及机器学习技术针对数据对青稞生长过程进行分析。故本文将参考已有禾本植物生长模型及西藏农牧研究院提供的部分青稞生长数据,将传统公式模型中以环境因子为自变量、青稞生长状态为因变量的计算公式转化为涵盖多影响因子的决策树模型。从而在有限数据的情况下实现决策树的相对最优建树过程,并在之后的生产实践过程中不断优化模型,在不断的迭代中给当地牧民生产种植活动提供更有效的指导。

本研究共涵盖以下几部分:作物生长模型分析;决策树建树原理介绍及本研究中的使用方式;拟定数据将传统公式模型转化为决策树模型;决策树模型结论分析及优化。

1 生长模型分析

通过查阅大量国内外相关文献及研究,笔者最终选用 WheatGrow 作物生长模拟模型作为本研究的理论基础,在此理论基础上结合青稞田间试验数据与决策树算法得到优化后的青稞生长 - 预测模型。本文将分析使用 WheatGrow 作物生长模拟模型进行研究的合理性,并分析产量预测子模型中的每穗小穗数、每小穗籽粒数及籽粒重三部分公式的原理。

1.1 模型合理性

WheatGrow 作物生长模型通过对禾本科作物生长过程的研究与分析,综合考量小麦生长过程中的环境影响因子与人为干预因素,借鉴国外 CERES-Wheat 模型的先进之处,形成了符合我国小麦生产国情实际的区域性适应模型[2]。青稞作为禾本科谷类作物,具有类似的器官结构,其生长过程符合一般作物生长规律,对光照、温度、土壤水分与氮素等外界环境因素做出相似的生理反应,因此与 WheatGrow 模型在作物生长影响因子的研究中与青稞的适配程度较高。

此外,该模型的 6 个子模型均采用国际模型检验统计方法 RMSE 进行模拟值与观察值间误差分析,所得预测误差都是在可信范围之内[2]。其中产量子模型的预测误差较小,预测效果较为理想,因此在本研究初探阶段作为青稞作物生长模型中产量预测的基础具有可行性。

1.2 每穗小穗数

N_{ASP}是在计算禾本作物产量中需要的计算的作物每穗的小穗数,具体计算方法见公式(1)[3]。其中,$NLPs$、N_{TSP}分别为二棱期前后分化的叶原基转化的小穗原基的数量;FN、FW 在公式(1)中分别为氮素与水分因子对穗分化的影响指数,在模型数据分析时认为作物在充分水分与氮素条件下生长,在计算中忽略不计。

$$N_{ASP} = (NLP_S + N_{TSP}) * \min(FN, FW) \quad (1)$$

小穗数 N_{ASP}、二棱期前后转化小穗原基数 NLP_S、N_{TSP}的具体计算方法如下。

1.2.1 二棱期前可转化的叶原基数 根据 WheatGrow 作物生长模型的产量预测模型[3],可计算出二棱期前可转化的叶原基数(公式(2))。

$$NLP_S = \left[1 - \frac{100PS * (2 - 0.0167PVT)}{pef_{SD}}\right] * (3.5 + GDD_{SE} * LPr_{SE} + 2LN_D) + \frac{1.5}{pef_{SD}} \quad (2)$$

式中,PS 为光周期敏感指数[4],表示特定品种对光周期的敏感程度;PVT 为品种决定的生理春化时间[4];Pef_{SD}为光周期影响因子,为播种至二棱期的平均光周期日长与小麦临界日长的比率,取值范围(0,1);LPr_{SE}出苗前的叶原基分化的速率,取值为 0.018;LN_D 为二棱期的叶龄[5],取值范围[3, 7]。

GDD_{SE}是出苗所需的有效积温,其计算过程与播种深度 $SDepth$ 和播种时的土壤含水量 SWP 有关[4]。

$$GDD_{SE} = \begin{cases} 10.2SDepth + 40, SWP \geqslant 75 \\ 2.1 * (100 - SWP) + 10.2SDepth + 40 \\ SWP < 75 \end{cases} \quad (3)$$

1.2.2 二棱期后分化的小穗数 二棱期之后进一步分化的小穗数 N_{TSP}在受到光周期变化影响的同时,也与穗分化持续时间有较大相关性,而作物对于温度的敏感程度对小穗分化持续时间造成一定影响[3],计算方法见公式(4)。

$$N_{TSP} = TSP_{num}P * e^{-5*10^{-4}} * (pef_{DT} * TS) * GDD_{DT} \quad (4)$$

$TSP_{num}P$ 为每穗分化的潜在小穗数,属于品种参数的一种;GDD_{DT}为二棱期至顶小穗形成期的有效积温,为简化模型而定为常数 1; TS 为品种的温度敏感性因子,表示不同品种对温度变化敏感程度的差异[4];pef_{DT}为二棱期以后的光周期效应因子,取值为二棱期至顶小穗形成的平均日长与小麦的临界日长的比值。

1.3 每小穗籽粒数

每小穗结实粒数的计算取决于作物小花原基分化的能力,以及水分、氮素、温度等环境因子对小穗实际结实粒数的影响[3],具体计算方法见公式(5)。

$$N_{SGrain} = FLP * FSP * e^{-0.1(17 - TAV)} * \min(FW, FN) \quad (5)$$

其中,FLP 为作物中部每小穗能够分化的小花数,与

各个品种的遗传特性相关,研究数据表明青稞每小穗只分化一朵小花,则计算时 *FLP* 取常数 1;*FSP* 表示因每穗总小穗数不同而引起的每小穗小花数或结实粒数的空间变异度,由于青稞每株小穗数与其他禾本作物相比较少[5],为简化模型 *FSP* 在青稞模型计算中取常数值 1;*TAV* 为灌浆期至成熟期间的日均温; $TAVF_{op}$ 为小花发育期的最适温度,对禾本作物取值为 20 ℃;*FN*、*FW* 在公式(5)中分别表示氮素与水分对在小花生长发育过程的影响指数。

1.4 籽粒重

籽粒重的计算要充分考虑对籽粒干物质量的积累过程的影响因子,包括特征千粒重、温度以及水分的效应影响指数[3](见公式(6))。

$$GrainW = \frac{THGrainW}{1000FDF} e^{-\frac{0.162(20-TAV)}{TS}} * FW \qquad (6)$$

GrainW 表示实际生长条件下的作物生产籽粒粒重;*THGrainW* 表示特定品种作物在最适生长条件下的千粒重,体现品种作物的生长潜力;*FDF* 为表示品种在灌浆时期至成熟所积累发育时间的进程,取值范围为[0.8, 1.0],其数值越大表示作物早熟程度越高;*TAV* 为灌浆期至成熟期间的日均温; $TAVF_{op}$ 为作物灌浆至成熟期间的最适温度,对禾本作物可以取值为 21 ℃;*TS* 为品种的温度敏感性因子;*FW* 为灌浆期间的水分效应指数,认为作物灌浆期期间的生长发育过程不缺乏水分。

1.5 产量计算

最终单位面积小麦籽粒产量 *P*,由穗数、每穗结实粒数和粒重的计算得出[3],详细计算见公式(7)。

$$P = N_{ear} * N_{ASP} * N_{SGrain} * GrainW * 1000 \qquad (7)$$

其中, N_{ear} 为单位面积穗数,决定于茎蘖发生动态,通常为穗数的 1 ~ 2.5 倍; N_{ASP} 为每穗结实小穗数; N_{SGrain} 为每小穗平均结实粒数;*GrainW* 为单籽粒重(g)。

1.6 变量分析

综上计算,影响总产量的变量包括光周期因子 pef_{SD}、二棱期叶龄 LN_D、播种深度 *SDepth*、播种时的土壤含水量 *SWP*、灌浆期至成熟期间的日均温 *TAV*、以及不同品种遗传特性所决定的品种参数。其中品种参数包括:光周期敏感性因子 *PS*、生理春化时间 *PVT*、每穗分化的潜在小穗数 $TSP_{num}P$、温度敏感性因子 *TS*、二棱后的光周期效应因子 pef_{DT}、中部每小穗分化的潜在小花数 *FLP*、基本灌浆期因子 *FDF* 共 7 个参数。

2 决策树建模理论

决策树模型法是一种典型的分类方法,其过程是对样本区数据进行处理,并利用归纳算法生成可读的规则和决策树,而后使用决策树对新区数据进行分析预测。决策树模型是基于决策点和策略点等构成的树形图,主要用于序列决策或多级决策。它的本质是一颗由多个判断节点组成的树。在使用模型进行预测时,根据输入参数依次在各个判断节点进行判断游走,最后到叶子节点即为预测结果[6]。因此利用决策树算法对传统公式模型进行迁移可得到涵盖个影响因子的预测模型,同时通过分析模型可得到不同影响因子对青稞生长的影响优先程度,最终达到指导牧民具体种植时间的目的。本研究于初探阶段采用决策树理论中的 ID3 算法。

2.1 ID3 算法的基本原理

ID3 算法是一种用来构建决策树的分类算法,于 1975 年由 Quinlan 提出[7]。在众多决策树构建方法中,ID3 算法是最具影响力的分类算法[8]。ID3 算法通过“信息熵”作为核心,某一环境因素对青稞生长的不确定性越大时,所对应的信息熵的值就越大。通过选择最高信息增益的影响因子来帮助确定父节点的因素选择标准,通过该因素对数据集进聚类同时建立分支,再通过递归方法建立各节点的分支,最终生成用于预测的决策树模型。

计算各影响因子的信息增益时,需要选取影响因子中具有最高信息增益的一项作为数据集合的测试属性,然后创建一个结点并给予标记,给该因子的每个值创建树分支,并根据这些分支进行划分样本[9]。详细计算见公式(8) ~ (10)。

计算数据集的信息熵公式:

$$H(D) = -\sum_{k=1}^{K} \frac{|C_K|}{D} \log_2 \frac{|C_K|}{D} \qquad (8)$$

其中:C_k 表示集合 *D* 中属于第 *k* 类样本的样本子集。针对某个特征 *A*,对于数据集 *D* 的条件熵 *H*(*D*|*A*)为:

$$H(D|A) = \sum_{i=1}^{n} \frac{|D_i|}{|D|} H(D_i) = \sum_{i=1}^{n} \frac{|D_i|}{|D|} \left(\sum_{k=1}^{K} \frac{|D_{ik}|}{D_i} \log_2 \frac{|D_{ik}|}{D_i} \right) \qquad (9)$$

其中:D_i 表示 *D* 中特征 *A* 取第 *i* 个值的样本子集,D_{ik} 表示 D_i 中属于第 *k* 类的样本子集。

信息增益 = 信息熵 - 条件熵:

$$Gain(D,A) = H(D) - H(D|A) \qquad (10)$$

其中,信息增益越大表示使用特征 A 来划分所获得的“纯度提升越大”。

ID3 算法建树简单易懂,计算便捷,通过较快的速度可以得到一棵具有科学依据且较为优化的决策

树模型。因此于本研究的初探阶段可快速得到传统公式模型的迁移后的决策树模型,通过对所的模型的初步分析可为实验后续研究提供可行性指导。

2.2 模型迁移过程

前文所提到的 WheatGrow 模型主要形式为不同影响因子关于青稞生长结果的数学计算公式,未涉及各影响因子之间的比较,且无法整合成涵盖各影响因子的青稞生长模型,因此无法直接对牧民种植活动提供指导。为解决传统公式模型的问题,我们将利用决策树理论,将传统公式模型中的各公式各变量根据青稞生长环境进行标准化程度划分,并计算各影响因子的信息增益进而确定决策树各层级各节点所选变量,最终得到一个包含各影响因子的青稞生长 - 预测模型。具体实验步骤如下。

传统公式模型涉及影响因子(自变量)及预测产量(因变量)的标准化:根据西藏青稞生长环境确定不同因素,如:光照、温度、土壤有机物成分、水分等环境因素的平均指标以及体现不同因素对青稞生长影响不同程度的数值区间。对不同公式的预测结果进行形式统一,得到统一的生长变化程度指标,以便进行数据增益的计算。

根据青稞环境因素数值于不同程度的分布情况,拟定对应数据带入传统生长模型,得到各环境因素对作物生长的影响在传统公式模型中的预测结果。

计算不同影响因子对青稞生长结果的信息增益,取增益效果最高的影响因子进行数据区间划分,既确定决策树各节点的分支,并对各分支数据进行递归处理得到包含各影响因子的青稞生长 - 预测模型。

根据西藏农牧研究院青稞历史生长数据对模型进行验证及优化。

3 实验过程

青稞生长模型影响因子复杂,在 WheatGrow 这一传统公式模型向决策树模型的迁移过程中所使用的数据将参照西藏农牧研究院在青稞生长研究过程中的青稞生长实验数据,并针对西藏特殊地理环境进行模型数据范围的调整,通过模拟西藏本土环境中青稞生长数据的概率分布,使迁移得到的决策树模型既符合禾本植物的生长规律,同时又适应青稞的生长环境数据,最终达到描述青稞生长过程中各影响因子的作用并对产量进行预测的目的。

3.1 影响因子标准化

经过对产量模型的分析,我们得到了对最终总产量有相关影响的环境及品种遗传因子,包括光照、温度、降水、土壤肥力方面等。根据西藏农牧研究院的试验数据与指导知识,我们依照各影响因子数据的概率分布情况,对各个自变量的取值区间进行划分,具体划分方式如表 1。

3.2 数据计算及建树

为保证拟定数据接近真实数据,在按照西藏农牧研究院测量数据对各影响因子依概率进行数值拟定的基础上,将相关联的影响因子数据拟定接近,其余因子数值随机组合,并对不同影响因子条件的数据带入计算公式进行计算,得到最终产量。将产量划分为以下 5 个区间,并随机取出 20 % 的数据项留作测试数据集,其余 80 % 的数据项用于建树:分别计算个影响因子带来的信息增益,并选择增益最大的影响因子作为根节点,划分不同区间作为各节点

表 1 各变量取值区间划分

名称		区间 1	区间 2	区间 3	区间 4	区间 5
环境变量	pef_{SD}	(0.2,0.3)	(0.3, 0.4)	(0.4, 0.6)	(0.6, 0.7)	(0.7, 0.8)
	pef_{DT}	(0.2,0.3)	(0.3, 0.4)	(0.4, 0.6)	(0.6, 0.7)	(0.7, 0.8)
	SDepth	(0, 2)	(2, 4)	(4, 6)	(6,10)	\
	SWP[1]	(15, 30)	(30, 40)	(40, 50)	(50, 60)	(60, 75)
	TAV	(0, 10)	(10, 15)	(15, 20)	(20, 25)	(25, 30)
作物参数	LN_D[5]	(0, 2)	(2, 3)	(3, 4)	(4, 5)	(5, +∞)
	PS[3]	(0, 3)	(3, 5)	(5, 8)	(8,10)	\
	PVT[3]	(20, 30)	(30, 40)	(40, 50)	(50, 60)	\
	$TSP_{num}P$[3]	(10, 15)	(15, 20)	(20, 25)	(25, 30)	(30, 35)
	TS[3]	(1.5, 1.6)	(1.6,1.65)	(1.65, 1.7)	(1.7, 1.75)	(1.75,1.85)
	FDF	(0.8, 0.85)	(0.85,0.9)	(0.9, 0.95)	(0.95,1)	\
	THGrainW	(0,20)	(20, 40)	(40, 45)	(45, 55)	(55, +∞)

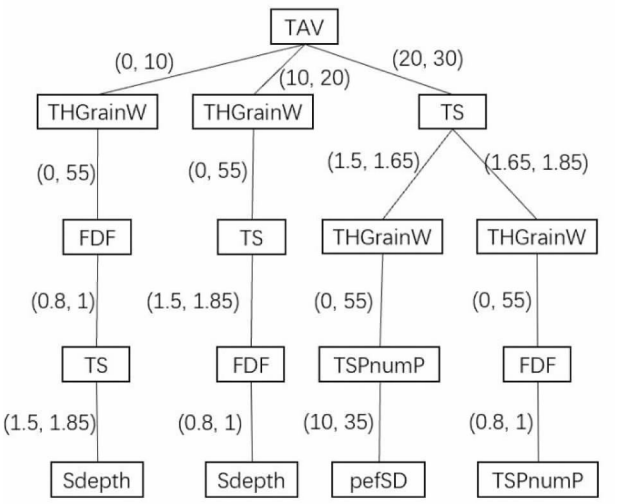

图 1　青稞生长影响因子决策树

的各分支,对各分支进行递归处理,最终能够得到青稞生长 - 预测决策树模型。数据分析工具采用 Excel2012 版本。

利用测试数据集对决策树进行印证,预测成功率达到 80 % 以上,可以较好地反应由传统公式模型所得数据集内的数据关系。其次,通过分析模型内各影响因子的影响优先级以及西藏青稞种植习惯,进一步提高所得模型的区域适应性,因此模型可用于描述青稞生长过程,并对西藏高海拔青稞的具体生产种植活动提供指导。

由于所建决策树层级较多且不同分支各影响因子影响所用效果较为相似,故对影响较小的因子及结构相同的子树进行合并操作,图 1 为保留 5 层节点的决策树缩略图。

3.3　结果分析

宏观分析来看,根据各影响因子所处层级可大致将以上影响因子划分为如下 3 类。

第 1 类,关键影响因子:*TAV*,即作物灌浆期日均温;*THGrainW*,即品种特征千粒重;*TS*,即品种温度敏感指数;*FDF*,即品种灌浆期早熟指数;$TSP_{num}P$,即品种每穗潜在分化小穗数;pef_{SD},即光周期因子;*SDepth*,即播种深度。

第 2 类,有影响因子:*SWP*,即土壤含水量;LN_D,即二棱期叶龄;*PS*,即作物光照敏感系数;*PVT*,即作物生理春化时间;

第 3 类,无影响因子:pef_{DT},即二棱期后光周期效应指数。

根据影响因子划分结果可知,牧民种植活动应根据模型所反映的不同影响因子对青稞产量影响的优先级进行逐级满足。既先关注影响青稞生长的关键影响因子,将有限的财力物力优先用于如保暖大棚、保暖地膜等建设及购置优质种子等相关事项,以达到投资收益的最大化。其次在满足关键影响因子达到青稞适宜条件后,适当关注有影响因子的对应条件,如进行土壤灌溉与适量施肥以保证作物生长的营养与水分供应等措施,使青稞产量进一步提高。生产过程中应尽量避免将过多的精力与成本投入到作物生长到无影响因子上,以实现生产效率最大化的目的。

此外,通过具体分析模型,牧民可针对青稞具体种植条件制定更为精确的生产种植策略。如在关键影响因子中,温度因素应作为牧民增产种植活动需考虑的首要因素,在优先保证温度在青稞适宜温度 19 ~ 20 ℃的同时,其次应注重青稞的品种选择。根据决策树二级分支可知,在温度较适宜的情况下应优先考虑选择温度较为敏感的品种以充分发挥温度在青稞生长和发育过程中的积极效应,而后进一步考虑对产量较优的青稞品种的研发与挑选,尤其是特征千粒重较为显著的品种。对于温度敏感程度较低的品种应进一步考虑培养具有较强分蘖潜力的品种;而对于受温度影响较大的作物中应充分利用品种的早熟程度,进而在更短的时间内得到更为客观的产量。根据图中 TAV 节点左子树可知,在温度较低的情况下应优先选择平均千粒重较大的青稞以保证基本产量的维持,其次在平均千粒重较大的种类中选择温度不敏感的品种以避免较低温度带来的影响;一旦遇到温度极低的年份,在选择平均千粒重较大的品种后应优先选择 FDF 较优的品种,极大程度上缩短极端温度对青稞生长和发育过程中产生负面影响的时间,而后再选择对温度敏感性尽可能低的品种。而后依次对其余各影响因子进行分析,达到精确指导种植的目的。

综上,本决策树模型既可从宏观上对各生长影响因子进行数据分析与区间细致划分,又可根据牧民具体种植条件判断所对应分支,从而进行更加精确的指导。本模型在传统公式模型的基础上实现了描述青稞生长过程各影响因子影响权重的功能,同时又可以因地制宜地对具体种植条件进行指导,最终更好地服务于西藏高海拔地区农牧民的青稞生产种植活动。

3　模型验证及改进方向

经研究对比发现,青稞生长 - 预测模型结果与研究院提供的青稞生长数据大致重合,且影响因子权重符合实际耕种经验。可初步确认生长验证模型从公式模型到决策树模型的迁移有一定有效效果。

但由于 ID3 算法没有剪枝策略,导致所有数据均被反映至决策树模型内,部分节点及分支的实际应用价值较低,后期模型优化工作应在收集足够真实种植数据的基础上展开相应的剪枝工作。此外,建树过程中还发现区间划分较多的变量对青稞产量结果有一定影响,其原因在于对信息增益的计算中,划分区间较多的影响因子对于变量与结果的相关性描述越准确,即容易计算得出较高的信息增益。因此,本研究在后续的模型优化或重建过程中应采用更为完善的 C4.5 算法,既引入信息增益率的概念,从而抑制由于单一变量区间划分过多而导致的信息及增益计算偏高的现象。最后,由于本次简述过程中所采用的数据均为用于迁移 WheatGrow 模型的拟定数据,使用数据过于理想,无缺值或不同测量标准数值的情况,在后期深入研究中,利用真实数据对模型进行优化的过程还需关注缺省数据的处理问题。

4 结 语

西藏生态环境较为脆弱,诸多因素限制了高海拔地区主要粮食作物青稞的大规模增产。长久以来,受制于没有统一的研究模型及衡量标准,增产研究难以量化。本文研究基于农业禾本植物生长模型,结合机器学习中决策树理论,初步建立起了涵盖青稞生长过程各影响因子的青稞生长 - 预测模型,为今后的研究提供了可参考的框架标准。尽管由禾本植物生长模型迁移得到的决策树模型精度并不理想,但是本研究实现了青稞生长 - 预测模型理论领域中从无到有的变化,并且模型可依靠每年收集的数据不断优化迭代,不断提高模型的准确性及精度。

参考文献:

[1]杨勤业. 西藏农田土壤水分状况的估算和灌溉问题[J]. 自然资源,1978(2):109 - 113.
[2]赵扬辉,汤亮,曹卫星,等. 小麦生长模拟模型(WheatGrow)的适应性评价[J]. 麦类作物学报,2010,30(3):443 - 448.
[3]潘洁,朱艳,曹卫星. 基于顶端发育的小麦产量结构形成模型[J]. 作物学报,2005(3):316 - 322.
[4]严美春,曹卫星,罗卫红,等. 小麦发育过程及生育期机理模型的研究 I. 建模的基本设想与模型的描述[J]. 应用生态学报,2000(3):355 - 359.
[5]扎桑,朱丽娟,卓嘎. 几种青稞新品种叶片出生与穗分化关系研究[J]. 湖北农业科学,2015,54(8):1932 - 1937.
[6]李舰,肖凯. 数据科学中的 R 语言[M]. 西安:西安交通大学出版社,2015.
[7]贾艺漩. ID3 算法的一种改进算法[J]. 信息通信,2019(5):14 - 15.
[8]Quinlan J R. Induction of decision tree[J]. Machine Learning,1986,4(2):81 - 106.
[9]Quinlan J R. Simplifying decision trees[J]. Internet Journal of Man-Machine Studies,1987,27(3):221 - 234.