

基于机器学习的苹果始花期预测

张兴伟, 陈超*, 田珊, 付琳

(成都信息工程大学, 成都 610225)

摘要: 始花期的早晚是生长过程中气象因子累积对果树产生的影响。冬季果树休眠是一个需冷量到需热量的过程, 研究了这一过程中气象因子对苹果果树始花期的影响并预测始花期。基于山西省吉县苹果物候数据和气象数据, 研究三个时间段内(是否发生冻害、能否正常越冬和热量与水分需求) 气温、湿度、地温、降水量和光照等气象因子对花期的影响程度, 采用多元线性回归方法和组合方法, 建立山西省吉县苹果果树始花期的预测模型。结果表明, 在三个不同时间段内预测值与真实值的决定系数分别为: 0.59、0.71 和 0.48。由于分析天数长短的不同, 基于能否正常越冬建立的模型效果最好, 其决定系数为 0.71, 分析天数最短的基于热量与水分需求时间段的模型决定系数最小, 这表明预测过程中分析天数的选择不宜过小。使用组合方法模型的决定系数为 0.78, 比能否正常越冬时间段模型的 0.71 略有提升。同时预报模型可以在 3 月 15 日完成提前量至少为 24 d 的精准预测。

关键词: 始花期预测; 相关性分析; 决定系数; 组合方法

doi: 10.13304/j.nykjdb.2019.0186

中图分类号: S126

文献标识码: A

文章编号: 1008-0864(2020) 10-0093-08

Prediction of Apple Initial Flowering Period Based on Machine Learning

ZHANG Xingwei, CHEN Chao*, TIAN Shan, FU Lin

(Chengdu University of Information Technology, Chengdu 610225, China)

Abstract: The early and late of the flowering period is the cumulative effect of meteorological factors on fruit trees during growth. Dormancy of winter fruit tree is a process that requires cooling to heat, so this paper studied the effect of meteorological factors on the flowering period of apple trees and predicted the initial flowering period. Based on the phenological data of apples and the meteorological data in Ji County, Shanxi Province, the effects of meteorological factors such as air temperature, humidity, ground temperature, precipitation and sunshine hours on the flowering period were determined to establish a prediction model for the initial flowering period of apple fruit trees at the three periods of whether frost damage occurs, whether it can overwinter normally, and heat and moisture requirements. by multiple linear regression methods and combined methods. The results showed that the coefficient of determination of the predicted value and the true value in three different time periods were 0.59, 0.71 and 0.48, respectively. Due to the different length of analysis days, the model based on the period of whether normally overwinter was the best with the coefficient of determination was 0.71, and the analysis with shortest days had the smallest coefficient of determination for the time of heat and moisture requirements, which indicated that the selection of analysis days should not be too small. The coefficient of determination of the combined method model was 0.78, which was a certain improvement over the 0.71 model based on whether it can overwinter normally. At the same time, the forecasting model could complete the accurate forecast of at least 24 d in advance on March 15.

Key words: prediction of initial flowering period; correlation analysis; coefficient of determination; combination method

收稿日期: 2019-03-18; 接受日期: 2019-04-18

基金项目: 成都市科技局项目(2018-YF05-01217-SN)。

联系方式: 张兴伟 E-mail: zxw-y@qq.com; * 通信作者 陈超 E-mail: RoosterChen@163.com

物候现象是生物节律与环境条件的综合反映。从气象条件来说,它不仅反映了当天的天气,而且反映了过去一段时间气象条件的累积情况^[1]。始花期预测正是基于这一理论而开展的研究。休眠是植物在生长过程中形成的一种对环境和季节性气候变化的生物学适应^[2],是植物发育中的周期性过程。休眠不仅可以使果树度过寒冬,而且也是落叶果树下一年正常开花结果必需经历的一个过程^[3]。果树生长过程中,当温度降低时落叶果树进入休眠期,冷量温度逐渐累积,果树自身通过一系列的生理变化应对温度降低,当冷量积累到一定程度,满足需冷量后开始升温,随即进入萌芽期开始累积热量,此时果树体内通过一系列生理变化来促进芽的萌发和生长^[4]。在苹果丰产栽培中,为了达到苹果高产、优质、稳产、高效之目的,必需将苹果营养生长规律与栽培管理密切结合起来。始花期的早晚是过去一段时间的气象条件累积对果树产生的影响。目前,关于苹果始花期的预测从特征向量上来分,有单气象因子与多气象因子两种。从算法模型上分,有统计学建模和线性回归建模。蒲金涌等^[5]和毛明策等^[6]研究了气温对苹果始花期的影响。李美荣等^[7]在果树物候模型理论的基础上,应用统计学方法对不同月份气温进行分析建模。柏秦凤等^[8]从苹果花期以前日平均温度的不同摄氏度积温及天数进行分析建模。张艳艳等^[9]使用最小二乘回归法进行多因子分析预测建模。藏曦^[10]使用不同月份的不同气象数据进行多因子逐步回归预测建模。同时,通过对其他果树的分析可以发现^[11-13],多因子线性回归建模为花期预测的主流预测方法。

多元线性回归预测建模能够很好的表征果树生长过程受不同气象因子的影响。但是目前的研究更多只是在算法层面,并没有很好的表明影响苹果始花期的主要影响时间段。根据山西省临汾市气象局发布的农用气象预报,研究分析休眠期内三个时间段的生长特性对苹果树的影响:是否发生冻害(12月1日至次年3月1日)、能否正常越冬(11月1日至次年3月15日)和热量和水分需求(3月1日至3月21日)。通过机器学习中的多元回归方法和组合方法预测始花期,从而可以得到影响始花期的主要时间段及主要影响气象因子,同时完成对始花期的提前、精准预测,以期

帮助果农提前做好农事安排和病虫害防治,为苹果果园清园、病虫害防治、田间管理和施肥提供建议,从而有助于增加果树树势和抵抗力,使苹果的经济效益最大化。

1 数据与方法

1.1 数据来源

气象资料包含吉县 1987—2017 年温度、降水量、湿度、地温、日照时长等气象因子,数据来源于山西省临汾市吉县气象局。

2010—2017 年苹果物候期资料由山西省临汾市吉县气象局提供,观测品种为“红富士”。

1.2 处理方法

1.2.1 时间间隔 以每年 1 月 1 日起至苹果果树开花始期为时间间隔,用于花期的预测。根据“中国物候观测网”的观测标准,植物始花期定义为观测植株上开始出现第一个完全开放花朵的日期^[14]。

1.2.2 数据标准化和相关性分析 使用 Z-score 标准化方法,将不同量级的数据统一转化为同一个量级,统一利用 Z 值进行衡量,以保证数据之间的可比性,消除由于不同量级数据所带来的影响。

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

式中, x 为观测值, μ 为总体平均值, σ 为总体标准差。

使用皮尔逊相关系数 (pearson correlation coefficient), 用于度量两个变量 x 和 y 之间的线性相关性。

$$\rho_{x,y} = \frac{\text{cov}(x,y)}{\sigma_x \sigma_y} \quad (2)$$

式中, $\text{cov}(x,y)$ 为 x 和 y 的协方差, σ_x 和 σ_y 分别为 x 和 y 的方差。

1.3 评价指标

均方误差 (MSE) 是指参数估计值与参数真值之差平方的期望值,它可以评价数据的变化程度, MSE 值越小,预测模型描述数据具有更好的精确度。

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - f_i)^2 \quad (3)$$

均方根误差 ($RMSE$) 是均方误差的算术平方根, $RMSE$ 值越小, 模型越好。

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - f_i)^2} \quad (4)$$

平均绝对误差 (MAE) 是绝对误差的平均值, 能够很好地反映预测值误差的实际情况。 MAE 值越小, 模型越精准。

$$MAE = \frac{1}{N} \sum_{i=1}^N |(y_i - f_i)| \quad (5)$$

式中, y_i 表示实际值, f_i 表示预测值。

决定系数 (R^2), 又称为判定系数或拟合优度, 它反映因变量的全部变动能通过回归关系被自变量解释的比例。表征了回归方程在多大程度上解释了因变量的变化, 或者说方程对观测值的拟合程度如何^[15]。 R^2 值越大越好, 当预测模型不犯任何错误时值为 1。

$$R^2 = \frac{SSR}{SST} \quad (6)$$

式中, SSR 为回归平方和, SST 为总平方和。

MSE 、 MAE 同样也是线性回归的损失函数, 损失函数的选取受到多方因素的制约和影响, 常见的影响因子有异常值、时间复杂度、求导困难度、预测值置信度等^[16]。在线性回归的时候目的就是让损失函数越小越好。

1.4 模型建立

1.4.1 相关性分析 在进行特征向量选择时, 尽可能剔除不相关或冗余的特征向量, 从而减少特征向量个数, 提高模型精确度, 减少模型运行时间。相关性分析可以帮助完成数据的筛选, 达到降维的目的。使用 `pandas` 库中的 `.corr` 函数, 分别对 12 月 1 日至次年 3 月 1 日 (是否发生冻害)、11 月 1 日至次年 3 月 15 日 (能否正常越冬) 和 3 月 1 日至 3 月 21 日 (热量和水分需求) 三个时间段的数据进行相关性分析。同时, 通过 `.nlargest` 函数设置参数 $k=4$ 选择出与时间间隔相关性最大的 4 个气象数据因子 (包含时间间隔), 完成多元线性回归模型中自变量的选择。

1.4.2 多元线性回归 回归模型应用训练集数据进行参数估计, 得到回归模型。如果回归分析中包含两个或两个以上的自变量, 且因变量和自变量之间是线性关系, 则称为多元线性回归。给定由 d 个属性描述的示例 $x = (x_1; x_2; \dots; x_d)$, 其

中, $y = w_1x_1 + w_2x_2 + \dots + w_dx_d + b$, 一般用向量形式表示 (式 7)。

$$f(x) = w^T x + b \quad (7)$$

式中, $w = (w_1; w_2; \dots; w_d)$ 。

当权重 w 和截距 b 学习得到之后确定模型。选择 30% 气象因子作为特征向量进行预测, 而剩余的 70% 为冗余向量。即选择与时间间隔相关性最大的 3 个特征 (不包含时间间隔) 作为特征向量, 分别对三个时间段进行多元线性预测建模。使用 2017 年数据为测试集, 验证模型准确性。剩余年份为训练集, 用于模型训练。三个时间段所对应的预测模型选用的特征向量分别为: 15 cm 地温、光照时长和 10 cm 地温; 15 cm 地温、10 cm 地温和 5 cm 地温; 最小湿度、平均相对湿度和降水量。

1.4.3 组合方法 由于在是否发生冻害和能否正常越冬时间段内的气象因子与时间间隔的正相关性因子远多于负相关因子, 故对这两时间段的预测结果使用“合奏”的方式, 对预测结果进行算术平均, 将算术平均后的结果进行取整。常用的方法有截取整数部分、向上取整、向下取整和四舍五入取整。通过实验验证发现, 通过向上取整后其结果更能拟合真实值结果。

最终结果 =

$$\text{Ceil} \left(\frac{\text{预测结果 1} + \text{预测结果 2}}{\text{个数}} \right) \quad (8)$$

1.5 模型实现

使用 Python 3.5 编程语言, 在 Jupyter notebook 中运行相应代码。

数据处理库 `numpy`、`pandas`, 数据可视化库 `seaborn`、`matplotlib`, 日期时间处理库 `datetime`, 机器学习库 `sklearn` 和数学函数库 `math`。

2 结果与分析

2.1 苹果果树的气温需求分析

根据临汾市气象局发布的农用天气预报可知, 冬季当地果树处于休眠期的平均气温要求在 $-10 \sim -7$ °C 之间, 当平均气温低于 -15 °C 时易发生冻害。以每年 12 月 1 日至次年 3 月 1 日是否为发生冻害的时间段, 分析 1987—2017 年的日平均气温 (图 1) 发现, 吉县苹果树在该时段内的气温在 $-4.52 \sim -0.54$ °C 之间, 满足吉县苹果树休眠期

的气温需求($-10\sim 7\text{ }^{\circ}\text{C}$),且日平均温度的平均值为 $-2.65\text{ }^{\circ}\text{C}$ 。这就说明对于吉县当地苹果树休眠期的气温需求可以放缩至 $-4.52\sim -0.54\text{ }^{\circ}\text{C}$ 的范围内。

2.2 苹果果树需冷量

落叶果树自然休眠所需的有效低温时数称为果树的需冷量,又称为低温需求量或需冷积温^[17]。

以每年11月1日至次年3月15日为能否正常越冬时间段。分析1987至2017年日平均温度 $\leq 7.2\text{ }^{\circ}\text{C}$ 的天数和日最高温度 $\leq 7.2\text{ }^{\circ}\text{C}$ 的天数,如图1所示。可以发现,在该时期内日平均气温符合 $\leq 7.2\text{ }^{\circ}\text{C}$ 的天数占比在88%以上,而日最高温度 $\leq 7.2\text{ }^{\circ}\text{C}$ 的天数占比平均保持在55.62%。这就说明在该时间段内,日平均温度 $<7.2\text{ }^{\circ}\text{C}$ 占比高于88%时,吉县苹果树能够顺利完成安全越冬。

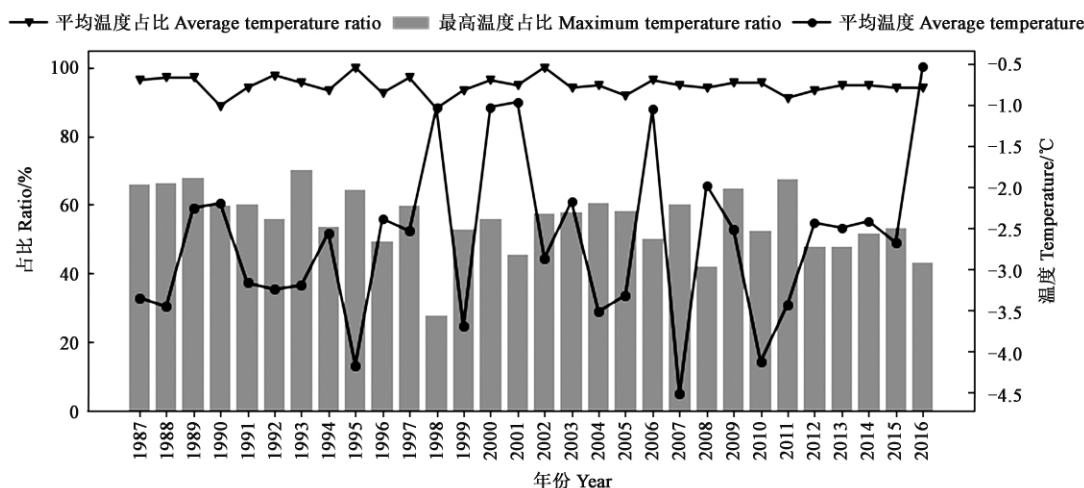


图1 前期气温分析

Fig.1 Preliminary temperature analysis

2.3 苹果树需热量和水分析

需热量是指从内休眠结束至盛花所需的有效热量累积,又称热量单位累积量或需热积温^[18]。当需冷量满足后,需热量则在一定程度上影响着果树花芽的正常萌发以及花期的早晚。引入降水量与湿度,用于分析水分对苹果树花期的影响。以每年3月上中旬(3月1日至3月21日)为热量与水分需求时间段。对2010—2017年该时间段的降水量、日最高温度、日最小湿度、光照时长和5 cm地温与时间间隔进行分析,结果如图2所示。可以看出,3月1日至21日部分气象数据与时间间隔的比较所示,2012年与2011年相比,当降水量和日最小湿度升高,日最高温度、光照时长和5 cm地温降低时,时间间隔降低;2013年与2012年相比,当降水量和日最小湿度降低,日最高温度、光照时长和5 cm地温升高时,时间间隔降低。说明在该时间段内,苹果树对于水分的需求与对热量的需求是相对反向的。同时也很好的说明了影响时间间隔的气象因子是多个的、复杂

的。与单因子预测建模相比,使用多因子的预测建模能够更好的表征气象因子对植物生长的影响。

2.4 苹果始花期花期预报结果分析

2.4.1 相关性分析 所有特征对结果的贡献是不一样的^[19],使用相关性分析用于对数据进行降维和特征选择。分别对是否发生冻害(12月1日至次年3月15日)、能否正常越冬(11月1日至次年3月15日)和热量与水分需求(3月1日至3月21日)三个时间段内标准化后的气象数据(2010至2017年)与时间间隔进行相关性分析,结果如表1所示。

从表1可以看出,在是否发生冻害时间段,温度、降水量、地温、湿度和光照时长与时间间隔都是正相关的,影响苹果树开花的主要因素是10、15 cm地温和光照时长,其与时间间隔的相关性都大于等于0.5;在能否正常越冬时期,15 cm地温是与时间间隔正相关性最大的因子,在该时期内比起对温度的关注更应该考虑地温;在热量

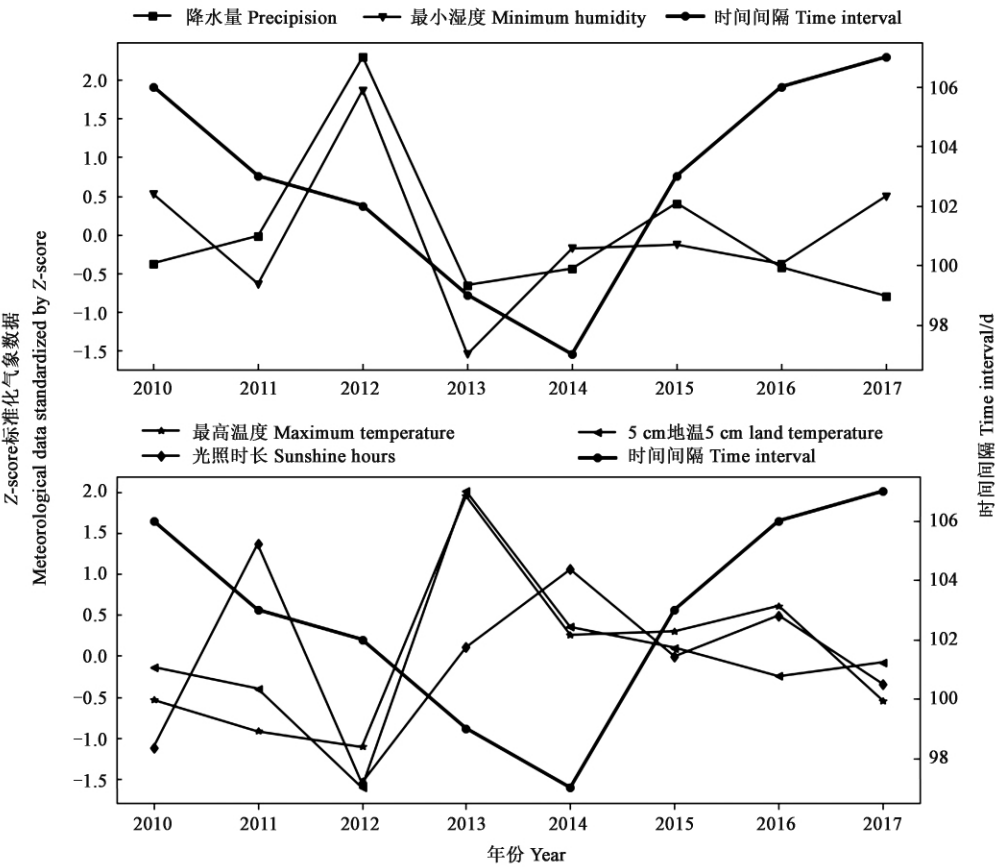


图 2 3 月 1 日至 21 日部分气象数据与时间间隔比较

Fig.2 Comparison of selected meteorological data with time intervals from 1st to 21st March

表 1 日气象因子与时间间隔的相关性

Table 1 Correlation between daily meteorological factors and time interval

模型 Model	日期 Date (mm/dd)	温度 Temperature			降水量 Precipi- tation	地温 Land temperature			相对湿度 Relative humidity		日照时间 Sunshine time
		平均 Average	最大 Max.	最小 Min.		5 cm	10 cm	15 cm	平均 Average	最小 Min.	
是否发生冻害 Whether frost damage occurs	12/01— 03/01*	0.31	0.15	0.34	0.30	0.44	0.50	0.58	0.15	0.13	0.56
能否正常越冬 Whether it can overwinter normally	11/01— 03/15*	0.20	-0.02	0.26	0.26	0.36	0.43	0.58	0.31	0.33	0.13
热量与水分需求 Heat and moisture requirements	03/01— 03/21	-0.38	-0.41	-0.33	-0.11	-0.41	-0.39	-0.33	0.1	0.32	-0.34

注: * 涉及跨年份数据。
Note: * time periods involve data across years.

与水分需求时期,只有湿度与苹果树开花量正相关,且日最小湿度为主要影响因子,这说明在 3 月中上旬,水分对苹果始花期的影响作用显著。

2.4.2 模型性能分析 通过对模型进行训练得到权重 W 与截距 b 如表 2 所示。
从表 3 预测模型检验可以看出,是否发生冻

害、能否正常越冬和热量与水分需求三个时间段的决定系数分别为 0.59、0.71 和 0.48,而通过组合方法的决定系数为 0.78,这说明组合方法有近 80%的对因变量的可解释性,而能否正常越冬模型只有 71%。当决定系数在 0.7 以上时,使用组

合方法的均方误差和平均绝对误差值明显低于能否正常越冬模型。均方根误差(*RMSE*)实质与均方误差一样,只是将误差的结果和数据同级化。四个算法模型的误差分别为 2.11、1.78、2.37 和 1.54 d。

表 2 模型结果

Table 2 Model results

模型 Model	W_1	W_2	W_3	b
是否发生冻害 Whether frost damage occurs	-0.101 709 14	2.040 019 11	1.813 366 24	102.788 680 07
能否正常越冬 Whether it can overwinter normally	9.578 495 65	-2.450 054 86	-5.725 328 29	102.790 779 38
热量与水分需求 Heat and moisture requirements	5.720 725 13	-4.148 460 20	-1.247 187 48	102.658 108 15

表 3 模型性能指标

Table 3 Test of predictive models

模型 Model	R^2	MSE	MAE	$RMSE$
是否发生冻害 Whether frost damage occurs	0.59	4.44	1.58	2.11
能否正常越冬 Whether it can overwinter normally	0.71	3.18	1.63	1.78
热量与水分需求 Heat and moisture requirements	0.48	5.62	2.18	2.37
组合方法 Combination method	0.78	2.38	0.88	1.54

通过热量与水分需求模型预测结果和是否发生冻害模型、能否正常越冬模型预测结果相比。是否发生冻害与能否正常越冬时间段的预测误差相对较低,数据之间正相关的因子远大于负相关因子,且决定系数都高于 50%。这是因为在研究过程中三个时间段的时间跨度不同。其分析天数分别为 21 d(热量与水分需求时间段)、90~91 d(是否发生冻害时间段)和 135~136 d(能否正常

越冬时间段),这就造成不同时间段内的气象因子与时间间隔的相关性不同,以致于影响评价指标。通过研究分析可以得出,在进行花期预测过程中气象因子的时间跨度选取不宜过小,且只有正相关因子才能提高预测的准确性。

从图 3 可以看出,使用组合方法获得的模型预测值与真实值在 2011、2012、2013 和 2017 年重合,在 2010、2015、2016 年保持 1 d 的误差。是否

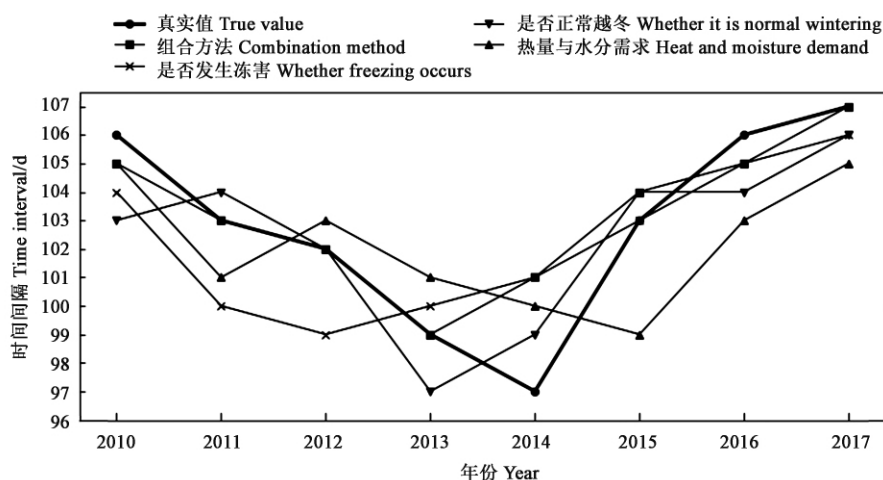


图 3 不同模型预测结果

Fig.3 Different model prediction results

发生冻害模型预测结果只有在 2015 年及以后误差较小,而在 2015 年之前平均误差 2 d。能否正常越冬模型模型预测值与真实值的误差 1~3 d,没有重合点,且大部分年份误差为 1 d。热量与水分需求模型其预测值与真实值的平均误差为 2 d。综上,结合衡量指标可以看出,能否正常越冬模型与组合方法模型是相对较好的两个模型,而热量

与水分需求模型波动较大,不适合用于预测。

2.4.3 花期预测结果 以 2017 年作为独立样本,将气象因子带入各预测模型中。对 2017 年苹果始花期花期进行试报,结果如表 4 花期预测值所示。不同模型的误差天数在 1~2 d 之间,花期预测模型拟合效果较好。

表 4 花期预测值

Table 4 Flowering forecast values

模型 Model	实际始花期 Actual beginning flowering period	预测始花期 Predicted flowering period	误差天数 Error days /d
是否发生冻害 Whether frost damage occurs	2017-04-18	2017-04-17	+1
能否正常越冬 Whether it can overwinter normally	2017-04-18	2017-04-17	+1
热量与水分需求 Heat and moisture requirements	2017-04-18	2017-04-16	+2
组合方法 Combination method	2017-04-18	2017-04-17	+1

组合方法模型的决定系数高于能否正常越冬模型和是否发生冻害模型,但其在试报过程中误差相同,这是因为得到的所有预测结果全部都是小数。而对于预测的结果,时间间隔要求是以天为单位的,这就需要对小数进行取整。在取整过程中对预测结果是采用向上取整、向下取整、四舍五入还是直接提取整数,具体哪种取整方式提出了要求,在此过程中便会降低预测的准确性。

3 讨论

本研究划分三个时间段,针对休眠期内苹果果树的关键时间节点气象因子对始花期的影响进行分析,分析天数分别为 21 d(热量与水分需求时间段)、90~91 d(是否发生冻害时间段)和 135~136 d(能否正常越冬时间段)。由于分析天数的不同造成了不同时间段内的气象因子与时间间隔的相关性不同,影响了评价指标和模型准确性。对于三个不同时间段的分析可以看出,在预测过程中的分析天数不宜过短,这与丁锡强等^[20]的研究结果相一致。对于分析天数的选取是否需要扩大,多久的分析天数才能更好的预测始花期这需要进行进一步的研究。

多元线性回归中,权重 w 直观表达了各属性在预测中的重要性,因此线性模型有很好的可解释性^[21]。每个数据的价值不一样,选择十个气象

因子中的三个气象因子作为特征向量参与建模,这是因为通过对是否发生冻害、能否正常越冬和热量与水分需求的三个时间段的训练集数据进行主成分分析后,发现三个特征向量分别可以表征 88.65%、97.26% 和 91.60% 的气象因子信息,即 30% 的特征向量表征了 88% 以上的信息。

组合方法思想认为好的模型并不一定是从多种模型中选择一个模型,而是可以把多个模型进行组合以得到更好的模型,包括“堆叠”和“合奏”两种思想。“堆叠”就是用前一个预测方法得到的预测值,为后一个预测方法得到的预测误差进行修正。“合奏”就是同时应用多个不同的预测方法得出多个预测值,并对这些预测值以某种方法进行平均,得到最终的预测值。采用的思想进行建模,是因为“合奏”的思想可以给出较高预测精确度的预测值,且模型简单、泛化能力好。

预测模型以 70% 以上的拟合程度,实现在 3 月 15 日的精准预报。所用数据集中始花期最早出现日期为 4 月 8 日,以 3 月 15 日进行预测可以提前 24 d 完成预测。这与尹贞铃等^[22]的 7~15 d 提前预测相比,在保证预测准确性的前提下预测的提前量有了很大的提升。同时,由于预测的时间间隔是以天为单位的,在模型预测过程中不可避免使用到小数取整的过程这对预测的准确性提出了新的挑战。

参 考 文 献

- [1] 国家气象局.农业气象观测规范(上卷)[M].北京:气象出版社,1993:133.
- [2] 付立中,申海林,邹利人,等.落叶果树自然休眠研究进展[J].北方果树,2014(4):1-3.
FU L Z, SHEN H L, ZOU L R, *et al.*. Research progress of natural dormancy of deciduous fruit tree [J]. Northern Fruits, 2014(4):1-3.
- [3] 欧阳汝欣,徐继忠,王同建,等.温度对果树芽休眠的影响研究进展[J].河北农业大学学报,2002(S1):101-103.
OUYANG R X, XU J Z, WANG T J, *et al.*. A study advance on effect of temperature on fruit trees bud dormancy [J]. J. Hebei Agric. Univ., 2002(S1):101-103.
- [4] 周慧,李晓瑞,达海莉,等.温度和光照对落叶果树休眠与萌芽的影响研究[J].宁夏农林科技,2018,59(10):28-29,41.
ZHOU H, LI X R, DA H L, *et al.*. Research progress of influence of temperature and light on deciduous fruit trees' dormancy and sprouting [J]. Ningxia J. Agric. Forestry Sci. Technol., 2018,59(10):28-29,41.
- [5] 蒲金涌,姚小英,姚晓红,等.气候变暖对甘肃黄土高原苹果物候期及生长的影响[J].中国农业气象,2008(2):181-183,187.
PU J Y, YAO X Y, YAO X H, *et al.*. Impacts of climate warming on phenological period and growth of apple tree in loess plateau of gansu province [J]. Chin. J. Agrometeorol., 2008(2):181-183,187.
- [6] 毛明策,刘敏茹,姜创业,等.苹果初花期与温度的关系研究[J].中国农业气象,2005(2):123-124,128.
MAO M C, LIU M R, JIANG C Y, *et al.*. A study on relationship between air temperature and early blooming time of malus [J]. Chin. J. Agrometeorol., 2005(2):123-124,128.
- [7] 李美荣,杜继稳,李星敏,等.陕西果区苹果始花期预测模型[J].中国农业气象,2009,30(3):417-420.
LI M R, DU J W, LI X M, *et al.*. Prediction model for beginning of apple flowering period in fruit growing areas of shaanxi province [J]. Chin. J. Agrometeorol., 2009,30(3):417-420.
- [8] 柏秦凤,王景红,屈振江,等.陕西苹果花期预测模型研究[J].中国农学通报,2013,29(19):164-169.
BAI Q F, WANG J H, QU Z J, *et al.*. The research on shaanxi apple florescence prediction model [J]. Chin. Agric. Sci. Bull., 2013,29(19):164-169.
- [9] 张艳艳,赵玮,高庆先,等.气候变化背景下陇东塬区'红富士'苹果始花期研究[J].果树学报,2017,34(4):427-434.
ZHANG Y Y, ZHAO W, GAO Q X, *et al.*. The effect of climate change on the apple's initial flowering data in the eastern gansu province [J]. J. Fruit Sci., 2017,34(4):427-434.
- [10] 藏曦.芮城县苹果花期预报模型研究[J].现代农业科技,2018(15):94-96,101.
- [11] 董海涛,单璐璐,谭丽静,等.基于最佳回归方法的丹东桃树花期预报研究[J].安徽农业科学,2019,47(1):232-238.
DONG H T, SHAN L L, TAN L J, *et al.*. Study on flowering forecast of dandong peach tree based on best regression method [J]. J. Anhui Agric. Sci., 2019,47(1):232-238.
- [12] 王萍,李仁杰,许海洋.基于多元回归的牡丹花期预测[J].农业网络信息,2008(3):139-142.
WANG P, LI R J, XU H Y. The forecasting of peony florescence based on multivariate regression analysis [J]. Agric. Network Inform., 2008(3):139-142.
- [13] 甘传辉,罗经权,万正,等.南昌县油菜花花期与气象因子的关系研究[J].江西农业学报,2018,30(6):114-118.
GAN C H, LUO J Q, WAN Z, *et al.*. Study on relationship between flowering date of rape plants and meteorological factors in Nanchang county [J]. Acta Agric. Jiangxi, 2018,30(6):114-118.
- [14] 张爱英,张建华,高迎新,等.SW 物候模型在北京樱花始花期预测中的应用[J].气象科技,2015,43(2):309-313.
ZHANG A Y, ZHANG J H, GAO Y X, *et al.*. Prediction of first flowering data of *Prunus discoides* in Beijing Yuyuantan Park using phenological model [J]. Meteorol. Sci. Technol., 2015,43(2):309-313.
- [15] 魏贞原.机器学习Python实践[M].北京:电子工业出版社,2018:95.
- [16] 文洁.MSE 与 MAE 对机器学习性能优化的作用比较[J].信息与电脑,2018(15):42-43.
WEN J. Comparison of the effects of mse and mae on machine learning performance optimization [J]. China Computer Commun., 2018(15):42-43.
- [17] 赵海亮,赵文东,高东升,等.落叶果树需冷量及其估算模型研究进展[J].北方果树,2007(6):1-3.
ZHAO H L, ZHAO W D, GAO D S, *et al.*. Advance in research of chilling requirement and estimating model on deciduous fruit trees [J]. Northern Fruits, 2007(6):1-3.
- [18] 谭钺.设施桃低温需求量与需热量关系机制的初步研究[D].山东泰安:山东农业大学,硕士学位论文,2010:10.
TAN Y. Preliminary studies on relationships between chilling requirements and heat requirements of peach in protected culture [D]. Shandong Taian: Shandong Agricultural University, Master Dissertation, 2010:10.
- [19] JOSHI P. Python 机器学习经典实例[M].陶俊杰、陈小莉,译.北京:人民邮电出版社,2017:17.
- [20] 丁锡强,王冰,姜茹茵,等.烟台市红富士苹果始花期预报模型研究[J].陕西气象,2018(3):33-36.
DING X Q, WANG B, JIANG R Y, *et al.*. Study on predicting model of flowering period for red fuji apple in yantai city [J]. J. Shaanxi Meteorol., 2018(3):33-36.
- [21] 周志华著.机器学习[M].北京:清华大学出版社,2016:53.
- [22] 尹贞铃,雪婷,辛博,等.白水苹果始花期预测模型研究[J].陕西气象,2019(2):34-36.
YIN Z Q, XUE T, XIN B, *et al.*. Study on the predicting model of the apple initial flowering period at baishui county [J]. J. Shaanxi Meteorol., 2019(2):34-36.

(责任编辑:温小杰)