



## تعیین متغیرهای اقلیمی مؤثر بر عملکرد پسته با استفاده از الگوریتم C&R درخت تصمیم

سمیه صدر<sup>۱\*</sup>، محسن اسلامی<sup>۲</sup>

تاریخ دریافت: ۱۳۹۹/۰۶/۲۶

تاریخ پذیرش: ۱۳۹۹/۱۲/۲۳

### چکیده

در دهه‌های اخیر، نوسانات اقلیمی در طی فصل رشد درخت پسته، کاهش مشهودی در عملکرد کمی و کیفی این گیاه راهبردی در مناطق عمده تولید آن در ایران ایجاد کرده است. تحقیق حاضر با هدف مشخص کردن مهمترین عوامل اقلیمی مؤثر بر عملکرد پسته با استفاده از مدل درخت تصمیم (C&R tree) انجام شد. مدل با استفاده از عوامل اقلیمی ایستگاه رفسنجان شامل سرعت باد، تعداد روزهای یخبندان، مجموع بارندگی، مجموع تبخیر، میانگین رطوبت هوا، جمع ساعات آفتابی، متوسط دما، میانگین کمینه دما و میانگین بیشینه دما (به عنوان متغیرهای ورودی) و عملکرد پسته (به عنوان متغیر هدف) طی سال‌های ۱۳۸۰-۱۳۹۸ اجرا شد. به منظور ارزیابی مدل، شاخص‌های آماری ضریب تبیین، جذر میانگین مربعات خطا، جذر میانگین مربعات خطای نسبی و اریبی استفاده شدند. درخت تصمیم برای میانگین ۶ ماه پیش از برداشت و ۶ ماه پس از برداشت (با تأثیر بر محصول سال بعد) به صورت جداگانه اجرا گردید و ضریب تبیین ( $R^2$ ) به ترتیب ۰/۸۸ و ۰/۸۶ بدست آمد که نشانگر مهارت قابل قبول رهیافت در تخمین عملکرد پسته می‌باشد. آماره RMSE برای ۶ ماه پیش و پس از برداشت محصول به ترتیب ۵۲۱ و ۵۵۸ ( $\text{Kg ha}^{-1}$ ) تعیین شد. بر اساس نتایج حاصل از کاربرد درخت تصمیم، از میان متغیرهای مستقل مورد استفاده در مدلسازی در نیمه اول سال مهمترین متغیرهای اقلیمی به ترتیب رطوبت نسبی و تعداد ساعات آفتابی و در نیمه دوم سال به ترتیب سرعت وزش باد، دمای کمینه، میزان بارندگی و رطوبت نسبی هستند.

**واژه‌های کلیدی:** رفسنجان، پسته، مدل‌سازی، نوسانات اقلیمی، درخت تصمیم

### مقدمه

تغییر در شرایط آب و هوایی محلی به ویژه در مراحل رشد و نمو محصولات منتج به کاهش عملکرد می‌شود (Bannayan et al., 2010). در دهه‌های اخیر، تغییرات اقلیمی (تغییر در فراوانی و شدت بارندگی‌ها، افزایش دما و خشکسالی) (Hatfield et al., 2011)، مشکلات زیادی از جمله عدم ریزش نزولات آسمانی و نوسانات شدید دمایی حاصل از آن در فصل پاییز، عدم تأمین نیاز سرمایی در زمستان، تغییرات ناگهانی دما در زمان تلقیح گل‌ها و افزایش روزهای گرم و سوزان تابستانی در سطح وسیعی به محصولات کشاورزی کشور آسیب زده است و باغ‌های پسته نیز از این موضوع مستثناء نیستند. محصول پسته به واسطه سطح بالای زیر کشت (حدود ۶۰ تا ۶۵ درصد سطح کشت پسته در دنیا (FAO, 2018)) و ارزش اقتصادی آن به عنوان یک محصول راهبردی برای کشور محسوب می‌شود. لذا بررسی این متغیرها بر عملکرد

شرایط آب و هوایی توصیف کننده شرایط زیست محیطی است که بر رشد محصول در یک منطقه مؤثر است. میزان عملکرد محصولات کشاورزی تحت تأثیر عوامل محیطی (زنده و غیر زنده) مانند شرایط آب و هوایی، خاک، آفات و خصوصیات ژنتیکی گیاه قرار می‌گیرد (Mahlein et al., 2012). اما تأثیر عوامل آب و هوایی بر روی عملکرد محصولات کشاورزی از اهمیت زیادی برخوردار است. ثابت شده است که در بیشتر موارد، بازده کم محصولات کشاورزی نتیجه عدم وجود شرایط متعادل جوی است (Siebert et al., 2017) و هر گونه

<sup>۱</sup> عضو هیئت علمی دانشگاه پیام نور استان کرمان، رفسنجان  
(\*نویسنده مسئول: 2716sadr@gmail.com)

<sup>۲</sup> کارشناس هواشناسی کشاورزی، مرکز تحقیقات هواشناسی کاربردی  
رفسنجان، کرمان

نحوه ارجاع مقاله:

صدر، س.، اسلامی، م. ۱۴۰۰. تعیین متغیرهای اقلیمی مؤثر بر عملکرد پسته با استفاده از الگوریتم C&R درخت تصمیم. نشریه هواشناسی کشاورزی،

DOI: 10.22125/agmj.2021.248249.1102 ۵۳-۶۲: (۱)۹۹

Sadr, S., Eslami, M. 2021. Determination of effective weather variables on pistachio yield using C&R decision tree algorithm. Journal of Agricultural Meteorology, 9(1): 53-62. DOI: 10.22125/agmj.2021.248249.1102

داده‌ها را می‌گیرند و یک درخت تصمیم را از آن استنتاج می‌کنند سپس درخت می‌تواند به صورت مجموعه قوانینی برای پیش‌بینی نتیجه ویژگی‌های معلوم استفاده شود. طی سال‌های اخیر از درختان تصمیم برای برآورد عوامل وابسته و اولویت‌بندی اثرمتغیرهای مستقل بر آن در علوم مختلف استفاده شده است. (Peloia et al., 2019) در پژوهشی به پیش‌بینی عملکرد محصول نیشکر با استفاده از الگوریتم درخت تصمیم پرداختند و نتایج نشان داد که درخت تصمیم می‌تواند در زمینه محصول نیشکر به‌عنوان یک سیستم پشتیبانی از تصمیم عمل کند و تجزیه و تحلیل پایگاه داده با استفاده از الگوریتم درخت تصمیم می‌تواند الگوهای قابل درک مربوط به عملکرد محصول را توصیف کند. (Arumugam 2017) پیش‌بینی عملکرد محصول برنج را با استفاده از الگوریتم درخت تصمیم انجام داد. دقت مدل درخت تصمیم ۹۲ درصد به دست آمد که از دقت بالایی برخوردار می‌باشد. نتایج این تحقیق به شناسایی متغیرهای تأثیرگذار در دستیابی به عملکرد بالا در محصول برنج کمک می‌کند. (Monjezi 2020) در عرضه‌یابی و اولویت‌بندی عوامل اجرایی و مدیریتی مؤثر بر بهره‌وری آب در تولید نیشکر از روش‌های درخت تصمیم CART<sup>۱</sup> و CHAID<sup>۲</sup> استفاده کرده و به ترتیب کمیت و کیفیت آب آبیاری، شرایط آب و هوایی، شرایط گیاه، عوامل مدیریتی و انسانی و شرایط خاک را در عملکرد محصول نیشکر مؤثر معرفی کرد. (Esmaeili et al., 2018) در تحلیل منطقه‌ای سیلاب از روش‌های درخت تصمیم با الگوریتم M5 در مقایسه با روش‌های رگرسیون خطی استفاده کرد. نتایج حاصل از بررسی آماره‌های صحت‌سنجی نشان داد بر اساس ضریب همبستگی بین آمار برآورد شده و مشاهداتی، همچنین بر اساس معیارهای جذر میانگین مربعات خطا و میانگین مطلق خطا، الگوریتم درخت تصمیم عملکرد بهتری نسبت به روش رگرسیون در برآورد دبی سیلاب دارد. در پژوهش دیگری درخت تصمیم در برآورد رسوبات رودخانه‌ای با روش مرسوم منحنی سنج رسوب<sup>۳</sup> مورد مقایسه قرار گرفت. نتایج این پژوهش نیز نشان داد که نتایج درختان تصمیم تطابق بسیار بیشتری با مقادیر اندازه‌گیری شده

محصول پسته، از اهمیت ویژه‌ای برخوردار است. در گذشته تحقیقات مربوط به پیش‌بینی محصولات کشاورزی با استفاده از متغیرهای آب و هوایی، بر پایه استفاده از تحلیل رگرسیون استوار بوده است اما از آنجا که تخمین عملکرد گیاه و متغیرهای مؤثر بر آن از روابط غیرخطی و پیچیده‌ای تبعیت می‌کند و مدلسازی آن نیز دشواری خاص خود را دارد، امروزه روش‌های نوین به عنوان روشی بهتر در برآورد عملکرد محصولات کشاورزی به کار رفته است. این روش‌ها علاوه بر صرفه‌جویی در وقت و هزینه به شناسایی عوامل مؤثر بر عملکرد نیز کمک می‌کند و قادر است روابط غیرخطی حاکم بین متغیرها را نیز شناسایی و مدلسازی کند (shirani, 2017). مطالعات متعددی در زمینه استفاده از روش‌های نوین به منظور مدل‌سازی عملکرد محصولات کشاورزی با استفاده از شاخص‌های اقلیمی انجام شده است. در این راستا می‌توان به مطالعه‌های (Cammarano et al., 2019) اشاره کرد که تأثیر تغییرات اقلیمی را بر عملکرد جو در اقلیم مدیترانه‌ای مورد بررسی قرار دادند. (Cai et al., 2019) از روش‌های یادگیری ماشین (شبکه عصبی مصنوعی و ماشین بردار پشتیبان) برای پیش‌بینی عملکرد گندم با استفاده از داده‌های اقلیمی و اطلاعات ماهواره‌ای بهره بردند و این مدل‌ها را برای پیش‌بینی عملکرد محصول گندم در استرالیا کارآمد معرفی کردند. در ایران نیز می‌توان به مطالعه برآورد عملکرد محصول کلزا (Mozaffari et al., 2016)، تخمین عملکرد زعفران (Akbarpour et al., 2013)، استفاده از متغیرهای هواشناسی و شاخص خشکسالی در برآورد عملکرد جو در استان آذربایجان غربی (Rahmani et al., 2010) و پیش‌بینی عملکرد گندم دیم در استان کردستان (Hosseini et al., 2007) اشاره کرد. یکی از روش‌های هوشمند و فراابتکاری برای مدل‌سازی عملکرد محصولات کشاورزی، روش درخت تصمیم است. درختان تصمیم از طریق جداسازی متوالی داده‌های ورودی به گروه‌های مجزا ساخته می‌شوند. این روش یک روش سلسله‌مراتبی است که در آن به صورت بازگشتی مجموعه داده‌ها به روش دودویی به تقسیمات فرعی و کوچکتر تقسیم می‌شوند تا زمانی که تقسیمات فرعی نهایی نتوانند بیشتر از آن تجزیه شوند. بنابراین درختان تصمیم مجموعه‌ای از

<sup>1</sup> Classification and Regression Trees

<sup>2</sup> Chi-squared Automatic Interaction Detection

<sup>3</sup> Sediment Rating Curve

تا ۵۶ درجه و ۳۴ دقیقه طول شرقی قرار گرفته است (شکل ۱). داده‌های مورد استفاده در این تحقیق مربوط به آمار ماهیانه سرعت باد، تعداد روزهای یخبندان، مجموع بارندگی، مجموع تبخیر، میانگین رطوبت، جمع ساعات آفتاب، متوسط دمای هوا، میانگین کمینه دما و میانگین بیشینه دما در ایستگاه سینوپتیک رفسنجان می‌باشد (جدول ۱).

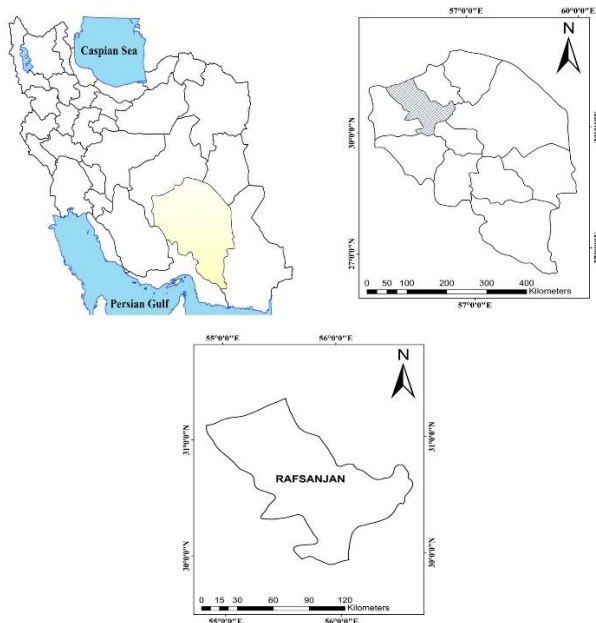


Figure 1- Geographic location of the study area

شکل ۱- موقعیت جغرافیایی منطقه مطالعاتی

Table 1- Independent variables used in this study

جدول ۱- متغیرهای مستقل مورد استفاده در این پژوهش

Variables	Unit
Wind speed (W)	$m s^{-1}$
Average temperature ( $T_m$ )	$^{\circ}C$
Total rainfall ( $R_t$ )	mm
Total evaporation ( $E_t$ )	mm
Average humidity (M)	%
Sum of sunshine hours (Sun)	-
Number of frost days (Z)	-

### مدلسازی درخت تصمیم

استفاده از درختان تصمیم، یک فرایند دو مرحله‌ای می‌باشد. در مرحله اول، مدلی براساس یک الگوریتم طبقه‌بندی مربوط به مجموعه آموزشی ساخته می‌شود. مجموعه آموزشی به صورت تصادفی از پایگاه داده انتخاب می‌شود. در مرحله دوم، یادگیری از طریق یک تابع  $y=f(X)$  انجام می‌شود که می‌تواند برچسب کلاس هر رکورد (x) از پایگاه داده را پیش‌بینی کند. مرحله

داشته و مهمترین عوامل مؤثر در مقادیر رسوب توسط درخت تصمیم در این پژوهش معرفی گردید ( Talebi and Akbari, 2013). از سایر این پژوهش‌ها می‌توان به مطالعات (Dastourani et al., 2013), Mozaffari et al., (2016) و Halabian et al., (2018) اشاره کرد. نتایج این پژوهش‌ها نیز به طور عمده درخت تصمیم را به عنوان ابزاری دقیق و قابل اعتماد معرفی می‌کنند و معتقدند که استفاده از این ابزار به عنوان روشی نو و قدرتمند می‌تواند در برنامه‌ریزی‌های کلی، یاری‌گر مدیران و برنامه‌ریزان باشد. در راستای استفاده از این روش در تعیین متغیرهای اقلیمی مؤثر بر عملکرد محصولات کشاورزی، می‌توان به مطالعه (Zakidizaji et al., 2019) اشاره کرد. ایشان از روش‌های درخت تصمیم در مدل‌سازی متغیرهای مؤثر بر عملکرد نیشکر در استان خوزستان استفاده کردند. نتایج این پژوهش تکنیک‌های داده‌کاوی، از جمله مدل‌های درختی را در برآورد عملکرد نیشکر کارا و مفید نشان داد. بر اساس بررسی‌های انجام شده، تاکنون هیچ‌گونه پژوهشی در زمینه شناسایی مهمترین عوامل اقلیمی مؤثر بر عملکرد محصول پسته به خصوص توسط روش‌های نوین مدل‌سازی انجام نشده است. لذا در این مطالعه با توجه به کم بودن تعداد داده‌های تأیید شده برای متغیر هدف (عملکرد) و قابلیت درخت تصمیم در شناسایی روابط پیچیده بین متغیر وابسته و مستقل در تعداد پایین داده، از این روش برای پیش‌بینی مهمترین متغیرهای اقلیمی مؤثر بر عملکرد محصول پسته مورد استفاده قرار گرفت.

### مواد و روش‌ها

#### منطقه مورد مطالعه و آماده‌سازی داده‌ها

شهرستان رفسنجان در استان کرمان، با حدود ۸۰۰۰۰ هکتار سطح زیر کشت پسته، به عنوان بزرگترین مرکز تولید پسته در ایران محسوب می‌شود ( Agricultural Statistics, 2017). لذا منطقه مورد مطالعه شامل بخش‌هایی از اراضی زیر کشت پسته در استان کرمان و در محدوده شهرستان رفسنجان است. رفسنجان با وسعتی حدود ۷۶۷۸ کیلومتر مربع در شمال غرب استان کرمان واقع شده‌است و در محدوده ۲۹ درجه و ۵۱ دقیقه تا ۳۱ درجه و ۳۱ دقیقه عرض شمالی و ۵۴ درجه و ۵۲ دقیقه

یادگیری، خود طی دو گام اساسی رشد و هرس انجام می‌شود. در طول فرایند رشد، الگوریتم درخت تصمیم می‌بایست به صورت مکرر مؤثرترین روش به منظور تقسیم کردن مجموعه ویژگی‌ها به فرزندان را بیابد. مرحله هرس برای جلوگیری از پردازش بیش از حد و بزرگ شدن درخت تصمیم که باعث پیچیدگی و افزایش تعداد قوانین اگر - آنگاه می‌شود، صورت می‌گیرد (Sattari et al., 2014). در این مطالعه از الگوریتم Classification and Regression (C&R) Tree استفاده گردید. الگوریتم C&R Tree، یک درخت باینری یک متغیره ایجاد می‌کند یعنی شاخه‌های خود را به صورت دوتایی و تنها بر اساس یک فیلد (متغیر مستقل) ایجاد می‌کند و هر گره غیر برگ آن، به دو گره دیگر تفکیک می‌شود. معیار انشعاب و شاخه‌زدن در درختان رگرسیون براساس حداقل کردن گوناگونی زیرمجموعه داخلی است. گوناگونی کم عبارت است از مجموعه‌هایی که اعضای یک کلاس در آن بر سایر کلاس‌ها غلبه کند و بهترین شاخه‌زدن آن است که گوناگونی در مجموعه‌ها را تا حد امکان کم کند. این الگوریتم برای متغیرهای کمی توسعه یافته است اما می‌تواند برای متغیرهای دیگر نیز مورد استفاده قرار گیرد. در این الگوریتم از ضریب استاندارد جینی (شاخص جینی) برای تقسیم داده‌ها در گروه‌های مختلف استفاده می‌شود و همچنین می‌توان از شاخص‌هایی مانند آنتروپی با سرعت بالاتر استفاده کرد. (Yohannes, 1999). به‌منظور اجرای مدل درخت تصمیم از آمار و اطلاعات مربوط به برخی متغیرهای اقلیمی که در طول سال‌های ۱۳۸۰ تا ۱۳۹۸ در ایستگاه سینوپتیک رفسنجان ثبت شده استفاده گردید. بر این اساس داده‌های هواشناسی برای هر سال به دو دسته، ماه‌های پیش از برداشت محصول (۶ ماه اول سال) و ماه‌های پس از برداشت محصول (۶ ماه دوم سال) تقسیم شد و در مدلسازی به عنوان متغیر ورودی (مستقل) مورد استفاده قرار گرفت. همچنین میزان عملکرد سالانه محصول پسته مطابق با آمار تأیید شده اداره جهاد کشاورزی رفسنجان برای ۱۸ سال متوالی (سال‌های ۱۳۸۰ تا ۱۳۹۸) تهیه و به عنوان متغیر هدف (عملکرد مشاهده شده) در مدلسازی مورد استفاده قرار گرفت. لازم به ذکر است که متغیرهای ۶ ماه دوم سال برای بررسی تأثیر این متغیرها بر محصول سال بعد مورد

استفاده قرار گرفت. در این پژوهش از بسته نرم‌افزاری ۲۴ statistics SPSS برای توصیفات آماری مربوط به داده‌های موجود و عملیاتی مانند رگرسیون چند متغیره مورد استفاده قرار گرفت و برای اجرای مدل‌سازی درخت تصمیم C&R از MATLAB نسخه‌ی ۲۰۱۳ استفاده گردید. در این مطالعه از آنجا که هدف تعیین مهم‌ترین متغیرهای اقلیمی مؤثر بر عملکرد بود و تخمین عملکرد و جامعیت دادن به مدل برای داده‌ها و مناطق دیگر مد نظر نبود تمامی متغیرها به عنوان داده آموزشی مورد استفاده قرار گرفت. در شبیه‌سازی با استفاده از مدل درخت تصمیم از متغیرهای مختلفی استفاده شد. این متغیرها به عنوان متغیر مستقل به مدل معرفی و شبیه‌سازی‌ها برای پیش‌بینی متغیر هدف صورت گرفت. متغیرهای مستقل مورد استفاده در این پژوهش، نقش و علائم اختصاری آن در جدول ۲ نمایش داده شده است.

### ارزیابی کارایی مدل

به منظور ارزیابی کارایی مدل درخت تصمیم در تخمین عملکرد، از متغیرهای آماری ضریب تبیین (معادله ۱)، جذر میانگین مربعات خطا (معادله ۲)، جذر میانگین مربعات خطای نسبی (معادله ۳) و انحراف<sup>۱</sup> (معادله ۴) استفاده شد.

$$R^2 = \frac{\sum(t-\bar{t})(y-\bar{y})}{\sqrt{\sum(t-\bar{t})^2 \sum(y-\bar{y})^2}} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum(y-t)^2}{N}} \quad (2)$$

$$\%RMSE = \frac{RMSE}{\bar{t}} \quad (3)$$

$$BIAS = \bar{t} - \bar{y} \quad (4)$$

که در آن‌ها، t مقادیر مشاهداتی و  $\bar{t}$  میانگین مقادیر مشاهداتی، y مقادیر تخمین زده شده و  $\bar{y}$  میانگین مقادیر تخمین زده شده و N تعداد مشاهدات است.

### نتایج و بحث

متوسط عملکرد در منطقه مورد مطالعه کمتر از ۴۹۲ کیلوگرم بر هکتار است. طبق آمار نامه کشاورزی، متوسط عملکرد پسته در استان کرمان ۴۴۷ کیلوگرم در هکتار است (Agricultural Statistics, 2017). خلاصه‌ای از

<sup>1</sup> Bias

کمتر از نصف کاهش پیدا کرده است. درخت پسته دمای ۴۵ درجه سانتی‌گراد در تابستان و دمای ۲۰- در زمستان را به خوبی تحمل می‌نماید (Hossaini fard et al., 2017) در این بازه زمانی تعداد روزهای یخبندان به طور متوسط به ۴۰ روز می‌رسد که طی این سال‌ها از ۵ تا ۶۷ روز متغیر بوده است. جدول ۳ ضریب همبستگی پیرسون بین عملکرد محصول پسته با متغیرهای اقلیمی را نشان می‌دهد.

**Table 3- Pearson correlation coefficient of pistachio crop yield with climatic characteristics**

جدول ۳- ضریب همبستگی پیرسون بین عملکرد محصول

پسته با متغیرهای اقلیمی مورد بررسی

Variable	y
W	-0.086
Z	0.045
R <sub>t</sub>	0.1
E <sub>t</sub>	-0.168
M	-0.08
Sun	0.05
T <sub>m</sub>	-0.18
T <sub>min</sub>	0.001
T <sub>max</sub>	-0.186

بر اساس جدول ۳، مشاهده می‌گردد همبستگی هیچکدام از متغیرها با عملکرد محصول پسته معنی‌دار نبوده و ضرایب همبستگی بسیار پایین بوده است که نشان‌دهنده عدم وجود رابطه خطی بین متغیرهای اقلیمی و عملکرد محصول پسته می‌باشد. همانطور که پیش از این اشاره شد رابطه تخمین عملکرد گیاه و متغیرهای مؤثر بر آن از روابط غیرخطی و پیچیده‌ای تبعیت می‌کند (Shirani, 2017). لذا استفاده از روش‌های رگرسیون در مطالعه حاضر نتایج قابل قبولی را نشان نداد. اما به طور مختصر بخشی از نتایج مدل‌سازی رگرسیون چند متغیره برای تعیین مهمترین متغیرهای مؤثر بر عملکرد محصول پسته طی سال‌های مذکور در جدول ۴ ذکر شده است. همانطور که مشاهده می‌شود پارامترهای مدل شامل عرض از مبدأ و ضرایب رگرسیون برای متغیرهای مورد مطالعه در ارتباط با عملکرد محصول پسته آورده شده است. اما آزمون t نشان می‌دهد که هیچکدام از متغیرها در مدل اثر معنی‌داری بر عملکرد ندارند زیرا آستانه معنی‌داری برای این متغیرها بزرگتر از ۰/۰۵ است و ضریب تبیین ( $R^2$ ) مدل ۱۵ درصد می‌باشد که بسیار کم است. مقدار منفی در  $Adj. R^2$  (جدول ۴) نشان می‌دهد که مدل، داده‌ها را به شکل ضعیفی برازش می‌کند. لذا مدل‌سازی عملکرد پسته با استفاده از خصوصیات آب و

توصیفات آماری مربوط به ۹ متغیر اندازه‌گیری شده، در جدول ۲ نشان داده شده است.

**Table 2- Statistical descriptions of the studied variables in the study area**

جدول ۲- توصیفات آماری مربوط به متغیرهای بررسی شده

در منطقه مورد مطالعه

Variable	Min	Max	S.D.	Aver.	Time
W	19.2	11.3	2.9	14.8	A
	20.5	11	2.5	16.2	B
Z	2	0	0.47	0.1	A
	67	5	14.8	39.2	B
R <sub>t</sub>	75	0.2	19.5	25.2	A
	96.2	9.8	24	48.5	B
E <sub>t</sub>	29.4	2167.8	277.1	2532.7	A
	n	n	n	n	B
M	25	19	1.8	21.9	A
	44	28.2	3.6	36.4	B
Sun	1990	1833	53.8	1918.5	A
	1554.6	1329.4	57.6	1440	B
T <sub>m</sub>	26.2	34.8	0.36	25.6	A
	13.5	11	0.8	12.5	B
T <sub>min</sub>	21	17.1	0.85	17.9	A
	6.3	2.7	0.99	5.2	B
T <sub>max</sub>	34	31	0.7	33.2	A
	21.1	18.4	0.86	19.8	B
y	1200	30	272.6	491.8	

A: ۶ ماه پیش از برداشت محصول، B: ۶ ماه پس از برداشت محصول، n: عدم وجود داده کامل

بر اساس نتایج بدست آمده در منطقه مورد مطالعه و طی بازه زمانی مد نظر، تفاوت قابل توجهی بین سرعت وزش باد در نیمه اول و نیمه دوم سال دیده نمی‌شود. مناسب‌ترین سرعت باد ۱۰ تا ۱۵ متر بر ثانیه در فروردین ماه است که برای بالا رفتن عملکرد گرده‌افشانی معرفی شده است (Hossaini fard et al., 2017). وجود بادهای ملایم در زمان گلدهی الزامی است ولی بادهای گرم و خشک، طوفانی و همراه با گرد و غبار مشکل‌ساز است و موجب تعرق شدید و خشک شدن مادگی و کاهش قدرت باروری دانه‌های گرده می‌شود و در مجموع گرده‌افشانی گیاهان را مختل می‌سازد و موجب خشکی خوشه‌ها می‌گردد. میزان بارندگی ثبت‌شده و درصد رطوبت در نیمه دوم سال تقریباً دو برابر نیمه اول سال است. برای داشتن حداکثر محصول، میزان رطوبت نسبی در تابستان بایستی کمتر از ۳۵ درصد باشد، البته وجود مقادیر کمتر رطوبت نسبی نیز توصیه شده است. این نکته را هم باید دانست که باران‌های کمتر از ۱۲ میلی‌متر اهمیتی از لحاظ تأمین رطوبت گیاه ندارند و تنها از طریق نقصان حرارت می‌توانند تأثیرگذار باشند. در بازه زمانی این مطالعه، متوسط دما در نیمه دوم سال نسبت به نیمه اول سال به

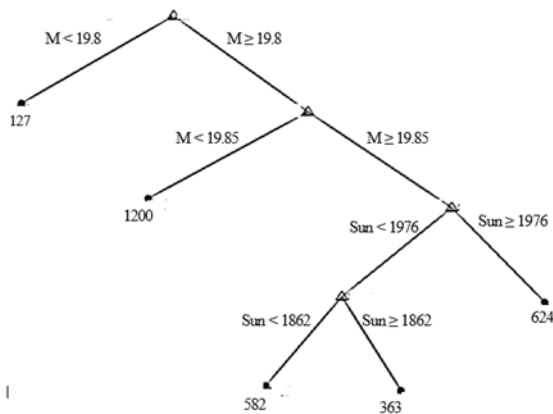


Figure 3- Decision tree created using data from 6 months after harvesting

شکل ۳- درخت تصمیم ایجاد شده با استفاده از داده‌های ۶ ماه پس از برداشت محصول

بر این اساس از میان متغیرهای مستقل مورد استفاده در مدل‌سازی در نیمه اول سال مهمترین متغیرها به ترتیب درصد رطوبت و تعداد روزهای آفتابی و در نیمه دوم سال به ترتیب سرعت وزش باد، دمای کمینه، میزان بارندگی و درصد رطوبت هستند. در مدل‌سازی متغیرهای مؤثر بر عملکرد نیشکر با استفاده از الگوریتم‌های درخت تصمیم‌گیر، عوامل مدیریت و واریته محصول در عملکرد نیشکر مؤثر معرفی شد (Zakidizaji et al., 2019). در پژوهش دیگری که برای مدل‌سازی عوامل اقلیمی مؤثر بر عملکرد سویا توسط درخت تصمیم انجام گرفت عوامل رطوبت نسبی، دما و بارندگی به عنوان عوامل با تأثیر بیشتر معرفی گردید (Veenadhari et al., 2011). در این مطالعه تنها به انتخاب مهمترین عوامل اقلیمی مؤثر بر عملکرد محصول پرداخته شده است و تأثیر مثبت یا منفی بر عملکرد نیاز به مطالعات بیشتر دارد لذا تأثیر مثبت و منفی درصد رطوبت چه در ماه‌های پیش از برداشت محصول و چه پس از آن را می‌توان چنین در نظر گرفت که افزایش رطوبت نسبی در زمان گلدهی و گرده‌افشانی باعث متورم و سنگین شدن دانه‌های گرده به علت جذب آب شده و فرایند انتقال و تلقیح با مشکل مواجه می‌گردد و با کاهش بازده گرده‌افشانی منجر به کاهش تشکیل میوه می‌شود. همچنین بالا بودن رطوبت در زمان رسیدن میوه باعث گسترش بیماری‌های قارچی می‌گردد (Crane, 1985). از طرفی تغییرات حرارتی هوای مرطوب بسیار کمتر از هوای خشک است و هوای مرطوب از یک سو فصل رشد را طولانی می‌کند و از سوی دیگر سبب

هوایی توسط الگوریتم درخت تصمیم انجام شد. در این مطالعه تست نرمال بودن توزیع خطاها مثبت بود و وجود داده پرت نیز بررسی گردید که تنها یک داده پرت در کل داده‌ها مشاهده و حذف گردید.

Table 4- The results of multivariate regression analysis to determine the most important climatic factors affecting the pistachio crop yield

جدول ۴- نتایج تحلیل رگرسیون چند متغیره در تعیین مهمترین عوامل اقلیمی مؤثر بر عملکرد محصول پسته

Variable	A		B	
	R.C.	Sig.	R.C.	Sig.
Constant	8563	0.50	-3015	0.41
W	34.6	0.75	-6.7	0.87
Z	-130.2	0.68	5.7	0.62
R <sub>t</sub>	-0.9	0.80	2.6	0.57
E <sub>t</sub>	-0.6	0.50	-0.5	0.72
M	-12.6	0.82	0.006	0.99
Sun	0.3	0.88	-	-
T <sub>m</sub>	-292	0.94	-1151	0.44
T <sub>min</sub>	64.7	0.98	443	0.42
T <sub>max</sub>	-29.5	0.99	777	0.39
	R=0.39		R=0.51	
	R <sup>2</sup> =0.157		R <sup>2</sup> =0.285	
	Adj. R <sup>2</sup> =0.079		Adj. R <sup>2</sup> =0.35	

A: ۶ ماه پیش از برداشت محصول، B: ۶ ماه پس از برداشت محصول، R.C.: Regression coefficients

شکل‌های ۲ و ۳ نمای درخت به دست آمده پس از هرس را نشان می‌دهد. یکی دیگر از خروجی‌های مدل درخت تصمیم، رتبه‌بندی اهمیت متغیرها است. معمولاً در مدل‌سازی بر متغیرهایی که اهمیت بیشتری دارند متمرکز می‌شود و متغیرهای کم اهمیت‌تر حذف می‌شوند. لذا متغیرهایی که در گره‌های بالاتر درخت قرار می‌گیرند اهمیت بالاتری دارند (Zakidizaji et al., 2019).

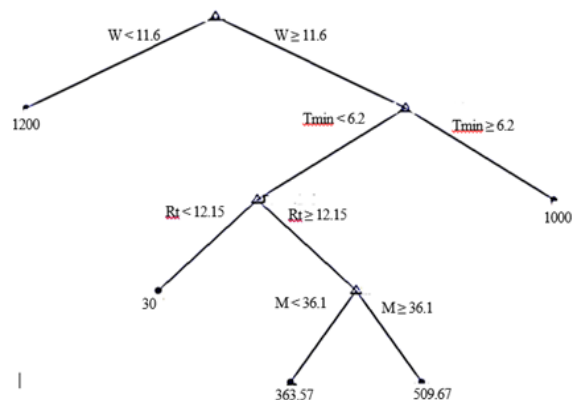


Figure 2- Decision tree created using data from 6 months before harvesting

شکل ۲- درخت تصمیم ایجاد شده با استفاده از داده‌های ۶ ماه پیش از برداشت محصول

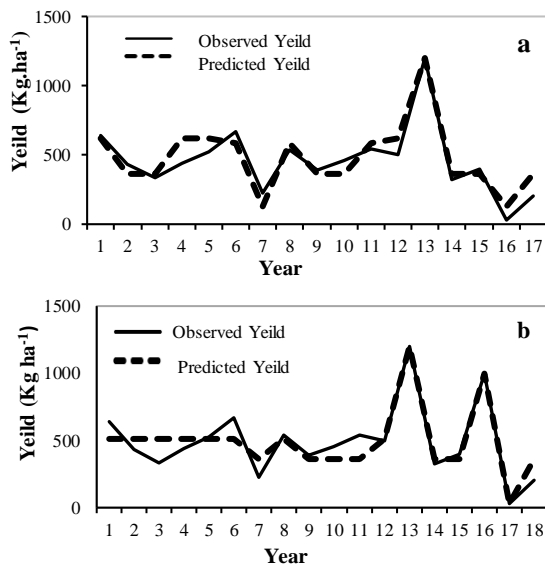
**Table 5- Quality of results from decision tree modeling in predicting pistachio crop yield at the first and second 6 months of the year**

جدول ۵- کیفیت نتایج حاصل از مدل‌سازی درخت تصمیم در پیش‌بینی عملکرد محصول پسته در ۶ ماهه اول و دوم سال

Parameter	A	B
R <sup>2</sup>	0.88	0.86
RMSE	521.2	558.7
RMSE%	1.128	1.135
BIAS	22.25	10

A: ۶ ماه پیش از برداشت محصول، B: ۶ ماه پس از برداشت محصول

بر این اساس در نیمه اول سال، درصد رطوبت و تعداد روزهای آفتابی توانست ۸۸ درصد از تغییرات عملکرد را پوشش دهد و در نیمه دوم سال سرعت وزش باد، دمای کمینه، میزان بارندگی و درصد رطوبت توانست ۸۶ درصد تغییرات عملکرد را پوشش دهد. همانطور که در جدول ۵ مشاهده می‌شود مقدار RMSE در نیمه اول به میزان ۶/۶ درصد نسبت به نیمه دوم سال کاهش نشان می‌دهد که تأکیدی است بر دقت بالاتر درخت تصمیم در مدل‌سازی عملکرد محصول پسته در نیمه اول سال که این موضوع در RMSE% نیز مشاهده می‌گردد. نمودار نقطه‌ای و نمودار خطی مربوط به نتایج پیش‌بینی عملکرد، توسط درخت تصمیم و مقادیر مشاهده شده، به ترتیب در شکل‌های ۴ و ۵ نشان داده شده است.



**Figure 4- Comparison of data observed and predicted by the decision tree for pistachio yield at the first half of year (a) and second half of year (b)**

شکل ۴- مقایسه‌ی داده‌های مشاهده شده و پیش‌بینی شده توسط درخت تصمیم برای عملکرد پسته در (الف) نیمه اول سال و (ب) نیمه دوم سال

مقاومت گیاه به تغییرات شدید دما در طول روز و شب می‌شود. اهمیت روزهای آفتابی در نیمه اول سال بسیار زیاد است. مدت زیاد تابش خورشید بخصوص در حالت عمودی باعث افزایش دما و افزایش طول مدت در معرض قرار گرفتن اندام گیاهی با دمای بالا می‌گردد. افزایش دما باعث کاهش رطوبت نسبی هوا می‌گردد و ظرفیت جذب رطوبت هوا زیاد می‌شود این حالت باعث تبخیر تعرق زیاد در باغ‌های پسته شده و از دست دادن آب گیاه افزایش می‌یابد بنابراین گیاه تحت تنش خشکی و کم آبی قرار می‌گیرد. تابش‌های شدید خورشید هدایت گرما به داخل میوه بخصوص در زمان‌هایی که ضخامت پوست سبز و پوست استخوانی کم است، می‌تواند به جنین در حال رشد خسارت وارد کند (Hokm Abadi, 2016). اهمیت دمای کمینه در نیمه دوم سال از این روست که به منظور کمک به بیداری و رویش جوانه‌ها و تولید گل آذین و تسریع در رشد رویشی درخت پسته حداقل به ۱۰۰۰ ساعت سرمای زمستانه ۲-۷ درجه سانتی‌گراد نیاز است. چنانچه جمع درجه سرمای زمستانه کمتر از ۶۷۰ ساعت باشد، گلدهی و برگ‌دهی درخت به تأخیر می‌افتد و برگ‌ها نامنظم و بدشکل می‌شوند و تولید محصول کاهش می‌یابد لذا درخت پسته باید در مناطقی کشت گردد که نیاز سرمای آن در طی فصل زمستان تأمین گردد. از آنجا که درجه حرارت و تبخیر آب از خاک در نیمه اول سال بالاست و همچنین در زمان برداشت محصول، آبیاری با تأخیر و یا کمتر از حد مورد نظر انجام می‌گردد، تجمع نمک در منطقه ریشه بروز می‌نماید لذا بارندگی در نیمه دوم سال هرچه بیشتر باشد شرایط خاک را از نظر تعدیل شوری بهتر خواهد کرد از طرفی بارندگی در مناطقی مانند رفسنجان که با بحران کمبود آب مواجه است می‌تواند جایگزین آبیاری زمستانه شود. این امر هم جذب عناصر غذایی که به عنوان کود در فصل خواب در اختیار درخت قرار می‌گیرد را بالا می‌برد و هم در کنار سرما و یخبندان زمستانه، بخش قابل توجهی از جمعیت بعضی از آفات که زمستان‌گذرانی آن‌ها در خاک می‌باشد را تعدیل می‌کند. در ادامه کارایی درخت تصمیم ایجاد شده با استفاده از شاخص‌هایی که پیش از این معرفی شد، بررسی گردید که مقادیر این متغیرها در جدول ۵ درج شده است.

امروزه استفاده مؤثر و استخراج داده‌های پنهان از انبوه داده‌ها و استفاده از تکنیک‌های داده‌کاوی می‌تواند به عنوان ابزار عملیاتی در پیش‌بینی عملکرد محصولات کشاورزی با توجه به انواع تغییرات اقلیمی مفید واقع شود. با وجود این که کنترل عوامل جوی و اقلیمی توسط انسان ناممکن است، اما انسان با تلاشی که در جهت ارتقاء دانش خود نسبت به تأثیر عوامل جوی دارد، می‌تواند توانمندی خود را در جهت کاهش خسارات ناشی از عوامل جوی به مرحله اجرا در آورد.

### نتیجه‌گیری

بر اساس نتایج این مطالعه مهمترین عواملی که عملکرد محصول استراتژیک پسته را تحت تأثیر قرار می‌دهد در ماه‌های پیش از برداشت محصول و پس از آن مشخص شد. با استفاده از الگوریتم رگرسیون درخت تصمیم، از میان متغیرهای مستقل مورد استفاده در مدلسازی در نیمه اول سال، مهمترین متغیرها به ترتیب درصد رطوبت و تعداد روزهای آفتابی و در نیمه دوم سال به ترتیب سرعت وزش باد، دمای کمینه، میزان بارندگی و درصد رطوبت هستند. لذا کشاورزان با تمرکز بر این موارد می‌توانند تا حدی از خسارت ناشی از کاهش عملکرد بکاهند. هر چند که لازم است ذکر گردد که بخش غیر قابل چشم‌پوشی از عملکرد پسته وابسته به مدیریت کودی و باغی است اما مجموعه عملیات مدیریتی در صورتی که به سمت سازگاری با تغییرات اقلیمی حرکت کند می‌تواند بهترین نتیجه را حاصل کند. از جمله این موارد می‌توان به عدم وجین کامل علف‌های هرز و انجام هرس فرم به منظور فاصله گرفتن شاخه‌های میوه ده از سطح زمین اشاره کرد. بر اساس نتایج حاصل از این پژوهش درخت تصمیم توانست روابط پیچیده بین متغیرهای مختلف اقلیمی با عملکرد محصول پسته را بهتر از رگرسیون خطی استخراج کند و مهمترین متغیرهای مؤثر را با دقت قابل قبولی مشخص کند.

### سپاسگزاری

بدینوسیله از دانشگاه پیام نور استان کرمان، مرکز رفسنجان قدردانی می‌شود.

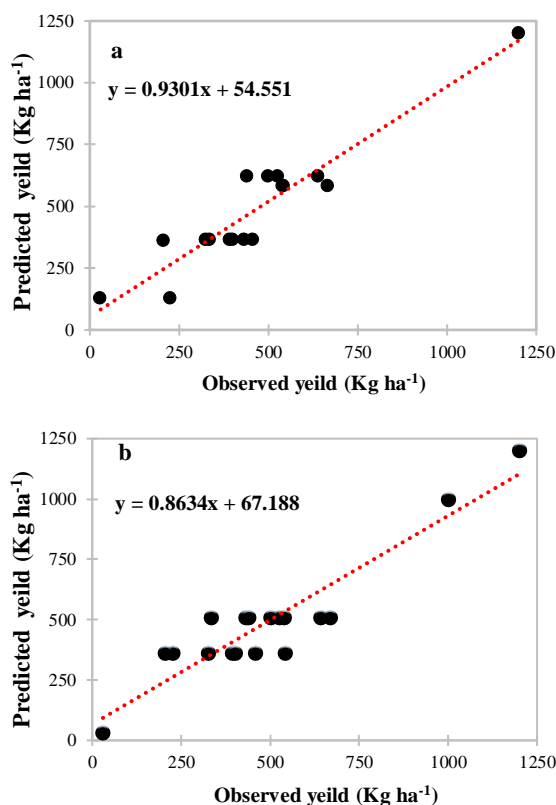


Figure 5- Comparison of data observed and predicted by the decision tree for pistachio yield at the first half of year (a) and second half of year (b)

شکل ۵- مقایسه‌ی داده‌های مشاهده شده و پیش‌بینی شده توسط درخت تصمیم برای عملکرد پسته در (a) نیمه اول سال و (b) نیمه دوم سال

همانگونه که مشخص است عملکرد درخت در مدلسازی بسیار خوب بوده به طوری که داده‌های پیش‌بینی شده و مشاهده شده به خوبی بر هم منطبق شده‌اند و اختلاف کمی را نشان می‌دهند. این نمودارها نشان می‌دهد که دقت برآورد عملکرد محصول پسته طی سال‌های آخر بالاتر از سال‌های اولیه در این پژوهش است. لازم به ذکر است که طی سال‌های اخیر (پس از ۱۳۸۸) ثبت آمار هواشناسی به صورت خودکار و هر ۱۰ دقیقه انجام می‌گیرد و این در حالی است که پیش از آن هر یک ساعت و به صورت چشمی انجام می‌شد. این امر دقت آمار ثبت شده را طی سال‌های اخیر بالا برده است که می‌تواند دقت بالاتر برآورد عملکرد طی سال‌های اخیر را توجیه کند. بدون شک تغییرات عملکرد محصول تحت تأثیر عوامل مختلفی است. در این تحقیق اثرات متغیرهایی همچون نوع خاک، تغییرات دمایی، آفات و امراض محصول و شیوه‌های مدیریتی در نظر گرفته نشده است.



## منابع

- Agricultural Statistics. 2017. Horticultural Products (Volume 3), Ministry of Jihad Agriculture, Deputy of Planning and Economy, Information and Communication Technology Center)
- Akbarpour, A., Khorashadzadeh, O., Shahidi, A., Ghochanian, E. 2013. Performance evaluation of artificial neural network models in estimate production of yield saffron based on climate parameters, *Journal of Saffron Research*, 1(1): 27-35. (In Farsi)
- Arumugam, A. 2017. A predictive modeling approach for improving paddy crop productivity using data mining techniques. *Turkish Journal of Electrical Engineering & Computer Sciences*, 25(6), 4777-4787.
- Bannayan, M., Sanjani, S., Alizadeh, A., Sadeghi, A., Lotfabadi, S., Mohammadian, S. 2010. Association between climate indices, aridity index, and rainfed crop yield in northeast of Iran. *Field Crops Research*, 118: 105-114.
- Cai, Y., Guan, K., Lobell, D., Potgieter, A.B., Wang, S., Peng, J., Peng, B. 2019. Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. *Agricultural and Forest Meteorology*, 274: 144-159.
- Cammarano, D., Ceccarelli, S., Grando, S., Romagosa, I., Benbelkacem, A., Akar, T., Ronga, D. 2019. The impact of climate change on barley yield in the Mediterranean basin. *European Journal of Agronomy*, 106:1-11.
- Crane, J.C. 1985. Pistachio. In *Handbook of fruit set and development*. CRC. Press.
- Dastourani, M.T., Habibipour, A., Ekhtesasi, M.R., Talebi, A., Mahjoobi, J. 2013. Evaluation of the Decision Tree Model in Precipitation Prediction (Case study: Yazd Synoptic Station), *Iran-water resources research*, 8(3): 14-23. (In Farsi)
- Esmaili, h., Mohammad Akhond Ali, A., Zarei, H., Taghian, M. 2018. Regional Flood Analysis Via Comparison of The M5 Decision Tree Algorithm and Regression Models, *Iranian Journal of Irrigation Sciences and Engineering*, 40(4):183-195. (In Farsi)
- Food and Agriculture Organization of the United Nations (FAO). 2018. <http://www.fao.org/faostat/en/#data/QC>.
- Halabian, A., Javari, M., Akbari, Z., Akbari, G. 2018. Evaluating the Performance of Decision Tree Model in Estimating the Suspended Sediments of river (A case study on the Basin of Meimeh River), *Geography and development Iranian Journal*, 15(49): 81-96. (In Farsi)
- Hatfield, J.L., Boote, K.J., Kimball, B.A., Ziska, L.H., Izaurralde, R.C., Ort, D., Thomson, A.M., Wolfe, D. 2011. Climate Impacts on Agriculture: Implications for Crop Production. *Agronomy Journal*, 103: 351-370.
- Hokm Abadi, H. 2016. Detection of environmental and non-environmental harmful factors on pistachio crop (insurance and compensation), *Agricultural Education Research Publications*, Tehran. (In Farsi).
- Hosseini, M.T., Siosemarde, A., Fathi, P., Siosemarde, M. 2007. Application of artificial neural network (ANN) and multiple regressions for estimating assessing the performance of dry farming wheat yield in Ghorveh region; Kurdistan province, *Agricultural research*, 7(1):41-54. (In Farsi).
- Hossaini fard, S.J., Basirat, M., Sedaghati, N., Ekhyani, A. 2017. Integrated fertility management and plant nutrition in pistachio trees, *Ministry of Jihad Agriculture, Pistachio Research Institute*. (In Farsi).
- Mahlein, A.K., Oerke, E.C., Steiner, U., Dehne, H.W. 2012. Recent advances in sensing plant diseases for precision crop protection. *European Journal Plant Pathology*, 133: 197-209.
- Monjezi, N. 2020. Diagnose and Prioritizing of Effective Managerial and Executive Factors on Water Productivity in Sugarcane Production and Providing Practical Solutions to Increase It. *Iranian Journal of Biosystems Engineering*, 51(1):99-112. (In Farsi)
- Mozaffari, G. A., Shafie; Sh., Tagizadeh, Z. 2016. Evaluate the performance Regression Decision Tree Model in Predicting Drought (Case Study: Synoptic Station in Sanandaj), 4(6):1-19. (In Farsi).
- Peloia, P.R., Bocca, F.F., Rodrigues, L.H.A. 2019. Identification of patterns for increasing production with decision trees in sugarcane mill data. *Scientia Agricola*, 76(4), 281-289.
- Rahmani, E., Liaghat, A., Khalili, A. 2010. Estimating Barley Yield in Eastern Azerbaijan Using Drought Indices and Climatic Parameters by Artificial Neural Network (ANN), *Iranian Journal of Soil and Water Research*, 39(1): 47-56. (In Farsi).
- Sattari, M.T., Abbasgoli Naebzad, M., Mirabbasi Najafabadi, R. 2014. Surface water quality prediction using decision tree method. *Journal of Irrigation and Water Engineering*, 4(3):76-88. (In Farsi).
- Shirani, H. 2017. Artificial neural networks with an application approach in agricultural sciences

- and natural resources, Valiasr University of Rafsanjan Publications. (In Farsi).
- Siebert, S., Webber, H., Rezaei, E.F. 2017. Weather impacts on crop yields searching for simple answers to a complex problem. Environmental. Research Letter. 12:081001.
- Talebi, A., Akbari, Z. 2013. Investigation of Ability of Decision Trees Model to Estimate River Suspended Sediment (Case Study: Ilam Dam Basin). Journal of Water and Soil Science, 17 (63) :109-121. (In Farsi).
- Veenadhari, S., Bharat, M., Singh, C. D., 2011. Soybean Productivity Modelling using Decision Tree Algorithms, International Journal of Computer Applications, 27(7):11-16.
- Yohaness, Y. 1999. Classification and Regression Tree: an Introduction. Research Institute of Washington, D.C.
- Zakidizaji, H., Bahrami, H., Monjezi, N., Shiekhdavoodi, M.J. 2019. Modeling of the Variables that Influence Sugarcane Yield using C5.0 and QUEST Decision Tree Algorithms. Journal of agricultural machinery, 9(2): 469-484. (In Farsi)



## Determination of effective weather variables on pistachio yield using C&R decision tree algorithm

S. Sadr<sup>1\*</sup>, M. Eslami<sup>2</sup>

Received: 16/09/2020

Accepted: 13/03/2021

### Abstract

In recent decades, climate variation during the pistachio growing have significantly reduced the quantity and quality of yield of this strategic crop of Iran. The purpose of this study is to determine the most important climatic factors affecting pistachio crop yield using C&R decision tree model in Rafsanjan region south of Iran during the period of 1380-1399. Modeling was performed using climatic variables including wind speed, number of frost days, rainfall, total evaporation, mean relative humidity, sunshine hours, mean temperature, mean of  $T_{max}$  and  $T_{min}$  as input variables and pistachio yield as the target variable. Correlation coefficient, root mean square error, relative error and bias metrics were used to evaluate the model. The decision tree model was run separately for 6 months before and after harvest. The obtained correlation coefficients were of 0.88 and 0.86, respectively, which is acceptable for prediction of yield. The RMSE values for pre- and post-harvest periods were 521 and 558 Kg ha<sup>-1</sup>, respectively. According to the results of the decision tree, during the preharvest period, the most significant attributes were the relative humidity and the number of sunny hours, respectively and for the second half of the year the wind speed, minimum temperature, rainfall and relative humidity were the most affecting variables, respectively.

**Keywords:** Decision tree, Modeling, Pistachio, Rafsanjan, Climatic variations



<sup>1</sup> Faculty Member, Department of Soil Science, College of Agriculture, Payame Noor University of Kerman, Iran

(\*Corresponding Author Email Address: [2716sadr@gmail.com](mailto:2716sadr@gmail.com))

<sup>2</sup> Agricultural meteorologist. Rafsanjan Applied Meteorology Research Center, Kerman, Iran