



---

## Identification of Cotton Fields Using the C5 Decision Tree Model and Multi-temporal Landsat Imagery: A Case Study of Golestan Province, Iran

Ghorban Ghorbani Nasrabad<sup>\*1</sup>, Khalil Ghorbani<sup>2</sup>, Maryam Mazidi<sup>3</sup>, Ebrahim Molaali<sup>4</sup>, Laleh Rezaei-Ghaleh<sup>5</sup>, Chooghi Bairam Komaki<sup>6</sup>, Abdolreza Ghoranjiki<sup>7</sup>

<sup>1\*</sup> Cotton Research Institute of Iran, Agricultural Research, Education and Extension Organization, (AREEO) Gorgan, Iran. Email: [ghorbang@yahoo.com](mailto:ghorbang@yahoo.com)

<sup>2\*</sup> Department of Water Engineering, Faculty of Water and Soil Engineering, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran.

<sup>3</sup> Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran.

<sup>4</sup> Cotton Research Institute of Iran, Agricultural Research, Education and Extension Organization, (AREEO) Gorgan, Iran.

<sup>5</sup> Postdoctoral Researcher, Department of Water Engineering, Faculty of Water and Soil Engineering, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran.

<sup>6</sup> Associate Professor, Department of Arid Zone Management, Faculty of Rangeland and Watershed Management, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran.

<sup>7</sup> Cotton Research Institute of Iran, Agricultural Research, Education and Extension Organization, (AREEO) Gorgan, Iran.

---

### Article Info

**Article type:**  
Research Full Paper

---

### ABSTRACT

**Background and Objective:** Accurate monitoring of crop cultivation patterns and crop separation at a regional scale is considered to be one of the fundamental requirements of sustainable water resources management and efficient agricultural planning. This study aims to evaluate and compare the performance of the C5 decision tree classification model in crop identification, with a particular focus on cotton, using Landsat 8 and 9 satellite imagery over the agricultural lands of Gorgan, Aliabad, and Aqqala counties in Golestan Province. The tree-based model used in this research was implemented stepwise, effectively reducing the greedy nature of the decision tree algorithm and identifying the most influential variables within the model.

### Article history:

Received: 2026-03-01  
Accepted: 2026-03-20

**Materials and Methods:** For this purpose, three time periods of August 6, August 31, and September 23 were selected to determine the most appropriate phenological stage of the growing season to achieve maximum cotton separation accuracy in multispectral data. The classification process was performed using the C5 algorithm as a rule-based model with automatic feature selection capability. A set of vegetation and moisture indices including NDVI, EVI, GNDVI, SAVI, NDWI and NMDI along with visible, near-infrared (NIR) and short-wave infrared (SWIR) spectral bands were used as input datasets.

### Keywords:

Crop classification

---

---

optical satellite  
C5  
spectral indices

**Results:** The comparative results of the error matrix showed that the date of September 23 provided the highest overall crop classification accuracy with a value of 88.10%; while the dates of August 6 and August 31 recorded accuracies of 83.6% and 82.3%, respectively. Also, the cotton crop was identified more successfully on the date of September 23 with an accuracy of 86.8% compared to other dates. The coincidence of this date with the full maturity stage of the cotton plant, which is accompanied by a decrease in chlorophyll, opening of bolls, a change in canopy structure and an increase in reflectance in the red and SWIR bands, creates a distinctive spectral signature of cotton compared to other crops.

**Conclusion:** These features have increased spectral separability and consequently improved the performance of the classification model at this time point. Accordingly, choosing the 23rd of September as the reference date and using the C5 model at this phenological stage can significantly improve the accuracy of identifying and mapping cotton fields in agricultural monitoring studies.

---

**Cite this article:** Ghorbani, Gh., Ghorbani, Kh., Mazidi, M., Molaali, E., Rezaei-Ghaleh L., Komaki, CB., Ghoranjiki, A.R., (2025). Identification of Cotton Fields Using the C5 Decision Tree Model and Multi-temporal Landsat Imagery: A Case Study of Golestan Province, Iran. *Iranian Journal Cotton Researches*, 13 (2), 1-23.

---



© The author(s)



10.22092/ijcr.2026.372160.1253

Publisher: Cotton Research Institute of Iran

---



## شناسایی مزارع پنبه با مدل درختی C5 و تصاویر چند زمانه لندست (مطالعه موردی: استان گلستان)

قربان قربانی نصرآباد<sup>۱\*</sup>، خلیل قربانی<sup>۲</sup>، مریم مزیدی<sup>۳</sup>، ابراهیم ملاعلی<sup>۴</sup>، لاله رضائی قلعه<sup>۵</sup>، چوقی  
بایرام کمکی<sup>۶</sup>، عبدالرضا قرنجیکی<sup>۷</sup>

- <sup>۱</sup> موسسه تحقیقات پنبه کشور سازمان تحقیقات، آموزش و ترویج کشاورزی، گرگان، ایران. رایانامه: [ghorbang@yahoo.com](mailto:ghorbang@yahoo.com)
- <sup>۲</sup> (نویسنده مسئول) گروه مهندسی آب، دانشکده مهندسی آب و خاک، دانشگاه علوم کشاورزی و منابع طبیعی گرگان، گرگان، ایران.
- <sup>۳</sup> دانشگاه علوم کشاورزی و منابع طبیعی گرگان، گرگان، ایران.
- <sup>۴</sup> موسسه تحقیقات پنبه کشور سازمان تحقیقات، آموزش و ترویج کشاورزی، گرگان، ایران.
- <sup>۵</sup> دانشکده مهندسی آب و خاک، دانشگاه علوم کشاورزی و منابع طبیعی گرگان، گرگان، ایران.
- <sup>۶</sup> گروه مدیریت مناطق بیابانی، دانشکده مرتع و آبخیزداری، دانشگاه علوم کشاورزی و منابع طبیعی گرگان، گرگان، ایران.
- <sup>۷</sup> موسسه تحقیقات پنبه کشور سازمان تحقیقات، آموزش و ترویج کشاورزی، گرگان، ایران.

اطلاعات مقاله	چکیده
<p>نوع مقاله: مقاله کامل علمی- پژوهشی</p> <p>تاریخ دریافت: ۱۴۰۴/۱۲/۱۱ تاریخ پذیرش: ۱۴۰۴/۱۲/۳۰</p>	<p><b>سابقه و هدف:</b> پیش دقیق الگوی کشت و تفکیک محصولات زراعی در مقیاس منطقه‌ای، از ارکان اساسی مدیریت پایدار منابع آب و برنامه‌ریزی کارآمد کشاورزی به‌شمار می‌رود. در این پژوهش، عملکرد مدل طبقه‌بندی درخت تصمیم C5 در شناسایی محصولات زراعی، با تأکید بر محصول پنبه، در اراضی کشاورزی شهرستان‌های گرگان، علی‌آباد و آق‌قلا واقع در استان گلستان با استفاده از تصاویر ماهواره‌ای لندست ۸ و ۹ مورد ارزیابی و مقایسه قرار گرفت. مدل درختی مورد استفاده در این پژوهش به شکل گام به گام اجراء شده که ضمن غلبه بر مشکل حریصانه الگوریتم درختی، متعیرهای مؤثر در مدل را نیز شناسایی می‌کند.</p>
<p><b>مواد و روش‌ها:</b> بدین منظور، سه مقطع زمانی ۱۶ مرداد، ۹ شهریور و دوم مهر به‌منظور تعیین مناسب‌ترین مرحله فنولوژیک فصل رشد برای دستیابی به بیشینه دقت تفکیک پنبه در داده‌های چندطیفی انتخاب شد. فرآیند طبقه‌بندی با بهره‌گیری از الگوریتم C5 به‌عنوان یک مدل قانون‌محور با قابلیت انتخاب خودکار ویژگی‌ها انجام گرفت. مجموعه‌ای از شاخص‌های پوشش گیاهی و رطوبتی شامل <i>NDVI</i>، <i>EVI</i>، <i>GNDVI</i>، <i>SAVI</i>، <i>NDWI</i> و <i>NMDI</i> به‌همراه باندهای طیفی مرئی، فروسرخ نزدیک (<i>NIR</i>) و فروسرخ کوتاه‌موج (<i>SWIR</i>) به‌عنوان ورودی مدل مورد استفاده قرار گرفتند.</p>	<p><b>یافته‌ها:</b> نتایج مقایسه‌ای ماتریس خطا نشان داد که تاریخ دوم مهر بالاترین دقت کلی طبقه‌بندی محصولات زراعی را با مقدار ۸۸/۱۰ درصد ارائه می‌دهد؛ در حالی که تاریخ‌های ۱۶ مرداد و ۹ شهریور به ترتیب دقت‌های ۸۳/۶ و ۸۲/۳ درصد را ثبت کردند. همچنین، محصول پنبه در تاریخ دوم مهر با دقت ۸۶/۸ درصد، نسبت به سایر تاریخ‌ها با موفقیت بیشتری شناسایی شد. هم‌زمانی این تاریخ با مرحله بلوغ کامل گیاه پنبه، که با کاهش کلروفیل، باز شدن غوزه‌ها، تغییر ساختار تاج پوشش و افزایش بازتاب در باندهای قرمز و <i>SWIR</i> همراه است، موجب ایجاد امضای طیفی متمایز پنبه نسبت به سایر محصولات زراعی می‌شود.</p>
<p><b>واژه‌های کلیدی:</b> طبقه‌بندی زراعی ماهواره نوری C5 شاخص‌های طیفی</p>	<p><b>نتیجه‌گیری:</b> این ویژگی‌ها سبب افزایش جدایش‌پذیری طیفی و در نتیجه بهبود عملکرد مدل طبقه‌بندی در این مقطع زمانی شده است. بر این اساس، انتخاب تاریخ دوم مهر به‌عنوان تاریخ مرجع و به‌کارگیری</p>

---

مدل C5 در این مرحله فنولوژیک، می‌تواند دقت شناسایی و نقشه‌برداری مزارع پنبه را در مطالعات پایش کشاورزی به‌طور معناداری ارتقا دهد.

---

**استناد:** قربانی نصرآباد، قربان؛ قربانی، خلیل؛ مزیدی، مریم؛ ملاعلی، ابراهیم؛ رضائی‌قلعه، لاله؛ کمکی، چوقی بایرام؛ قرنجیکی، عبدالرضا (۱۴۰۴). شناسایی مزارع پنبه با مدل درختی C5 و تصاویر چند زمانه لندست (مطالعه موردی: استان گلستان). *مجله پژوهش‌های پنبه ایران*، ۱۳ (۲)، ۱-۲۳.



10.22092/ijcr.2026.372160.1253

© نویسندگان

ناشر: موسسه تحقیقات پنبه کشور



## مقدمه

پایش کاربری اراضی و شناسایی دقیق نوع کشت در مقیاس مزرعه‌ای، یکی از مؤلفه‌های کلیدی در مدیریت پایدار منابع طبیعی و برنامه‌ریزی بهینه کشاورزی به شمار می‌آید. در بسیاری از مناطق کشاورزی کشور، به‌ویژه در استان‌هایی با تنوع اقلیمی بالا و الگوی کشت پویا همچون استان گلستان، اتکا به روش‌های سنتی شناسایی کاربری اراضی، شامل بازدیدهای میدانی، برداشت‌های کارشناسی و گزارش‌های محلی به دلیل هزینه‌بر بودن، زمان‌بر بودن، محدودیت دقت و ناتوانی در پوشش سطوح وسیع، کارایی لازم را ندارد. این در حالی است که تغییرات کاربری اراضی و نوسانات سطح زیرکشت محصولات راهبردی، به‌طور مستقیم بر سیاست‌های تخصیص منابع آب، تنظیم الگوی کشت، امنیت غذایی و برنامه‌ریزی اقتصادی اثرگذارند. از این‌رو، بهره‌گیری از داده‌های سنجش‌ازدور در کنار روش‌های نوین تحلیل داده، به‌عنوان راهکاری کارآمد و اجتناب‌ناپذیر برای پایش دقیق و بهنگام اراضی زراعی مطرح شده است. در این میان، تصاویر ماهواره‌ای لندست به دلیل برخورداری از قدرت تفکیک مکانی مناسب، پوشش زمانی منظم با دوره تناوب ۸ تا ۱۶ روزه و قابلیت استخراج شاخص‌های متنوع پوشش گیاهی، یکی از معتبرترین منابع داده‌ای برای تحلیل تغییرات کاربری اراضی و شناسایی محصولات کشاورزی محسوب می‌شوند. در مطالعات اولیه، سنجنده‌های  $TM$  و  $ETM^+$  ماهواره‌های لندست ۵ و ۷ امکان تحلیل‌های زمانی گسترده و پایش روند تغییرات پوشش زمین را فراهم کردند. برای نمونه، علیپور و همکاران (۲۰۱۴) با استفاده از تصاویر  $ETM^+$  توانستند سطح زیرکشت را در مزرعه آستان قدس رضوی با دقت قابل قبولی برآورد کنند. همچنین، کامکار و همکاران (۲۰۲۱) با تحلیل تصاویر لندست ۵، ۷ و ۸ در بازه زمانی ۲۰۰۰ تا ۲۰۱۶، تغییرات سطح زیرکشت گندم و سویا را در غرب استان گلستان بررسی کردند که نتایج آن بر توانمندی لندست در پایش تغییرات بلندمدت و تفکیک محصولات تابستانه تأکید دارد. افزون بر این،

دهقانی و همکاران (۲۰۲۲) نشان دادند که تصاویر لندست ۸ در ارزیابی تغییرات کاربری اراضی در مقیاس منطقه‌ای و شهرستانی، ابزاری دقیق و مقرون‌به‌صرفه به شمار می‌آیند. توانایی تفکیک طیفی مناسب تصاویر لندست در نوارهای مرئی، فروسرخ نزدیک و فروسرخ با طول موج کوتاه، این امکان را فراهم می‌سازد که تفاوت‌های فنولوژیک محصولات کشاورزی، حتی در مراحل رشد مشابه، شناسایی شود. به همین دلیل، مطالعات متعددی از این تصاویر در ترکیب با الگوریتم‌های طبقه‌بندی برای تفکیک کلاس‌های کاربری اراضی استفاده کرده‌اند. برای مثال، نتایج تحقیقات فریز و همکاران (۲۰۱۷) و پربتی و همکاران (۲۰۲۳) نشان می‌دهند که به‌کارگیری تصاویر لندست در چارچوب روش‌های یادگیری ماشین، دقتی بیش از ۸۰ درصد در طبقه‌بندی اراضی کشاورزی فراهم می‌کند؛ که بیانگر کارایی بالای این داده‌ها در شناسایی محصولات زراعی است. در اغلب پژوهش‌های مرتبط، طبقه‌بندی کاربری اراضی با تکیه بر ترکیبی از شاخص‌های طیفی، اطلاعات توپوگرافی و معیارهای تفکیک‌پذیری انجام می‌شود (ممتازی بروجنی و سرمدیان، ۲۰۲۲). در این میان، شاخص‌های پوشش گیاهی نقش محوری دارند؛ به‌طوری که شاخص  $NDVI$  به‌طور گسترده برای پایش تغییرات پوشش گیاهی و برآورد سطح زیرکشت محصولات مختلف مورد استفاده قرار گرفته است (گائو، ۱۹۹۶). افزون بر آن، شاخص  $SAVI$  برای کاهش اثر خاک در مناطق با پوشش گیاهی تنک و شاخص  $EVI$  به‌عنوان جایگزینی بهبودیافته برای  $NDVI$  در شرایط اشباع طیفی، در تفکیک محصولات زراعی به‌کار رفته‌اند (کامکار و همکاران، ۲۰۲۱)؛ عبیات و همکاران، (۲۰۲۱). در سال‌های اخیر، استفاده از الگوریتم‌های یادگیری ماشین برای طبقه‌بندی اراضی و پایش پوشش زمین مبتنی بر تصاویر ماهواره‌ای، به‌دلیل توانایی بالای آن‌ها در استخراج الگوهای پیچیده از داده‌های حجیم دیدبانی زمین، به یک حوزه کلیدی در مدیریت منابع طبیعی تبدیل شده است (ژانگ و همکاران، ۲۰۲۵). در این راستا، الگوریتم‌هایی نظیر حداکثر احتمال،

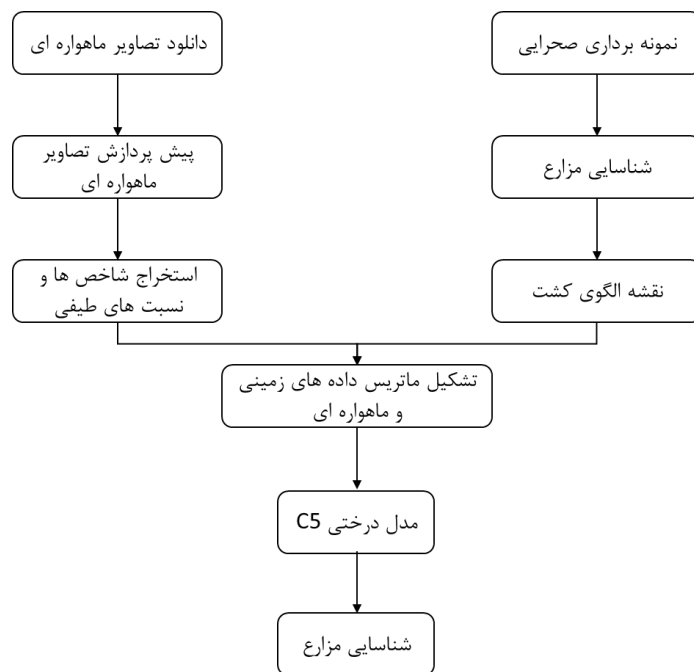
هدف شناسایی دقیق مزارع پنبه و پایش تغییرات آن‌ها، از تصاویر ماهواره‌ای لندست در سه تاریخ زمانی متفاوت و الگوریتم درخت تصمیم *C5.0* بهره می‌گیرد. فرض بر این است که (۱) شاخص‌های طیفی منتخب، به‌ویژه شاخص‌های پوشش گیاهی، قابلیت تمایز معنادار میان پنبه و سایر محصولات زراعی را فراهم می‌کنند؛ (۲) ترکیب این شاخص‌ها با الگوریتم *C5.0* موجب افزایش دقت طبقه‌بندی در مقایسه با روش‌های رایج می‌شود؛ و (۳) الگوریتم *C5.0* توانایی مدیریت هم‌پوشانی طیفی و کاهش افزونگی اطلاعات را دارد. در نهایت، هدف این مطالعه ارائه چارچوبی کارآمد برای استخراج الگوی کشت پنبه و پایش بهنگام آن در مقیاس منطقه‌ای است؛ چارچوبی که می‌تواند به‌عنوان ابزاری مؤثر در پشتیبانی از تصمیم‌سازی‌های مدیریتی و برنامه‌ریزی کشاورزی مورد استفاده قرار گیرد.

### مواد و روش‌ها

در این پژوهش ابتدا بر اساس کار میدانی، از موقعیت مزارع با کشت‌های مختلف نمونه‌برداری می‌شود. تصاویر ماهواره‌ای لندست پس از دانلود و تصحیح در زمان‌های مختلف طی دوره کشت، پیش‌پردازش شده و اطلاعات باندهای طیفی استخراج و بر اساس آنها شاخص‌ها و نسبت‌های طیفی محاسبه می‌شود. ماتریسی از آنها به همراه داده‌های عملکرد و نوع کشت تشکیل و به عنوان ورودی به مدل درخت تصمیم (برای شناسایی نوع کشت) داده می‌شود. در نهایت بر اساس تحلیل‌های آماری، امکان شناسایی مزارع ارزیابی می‌شود. در شکل (۱) مراحل پردازش تصاویر و تعیین مدل ارائه شده است.

ماشین بردار پشتیبان (*SVM*) و جنگل تصادفی (*RF*) به‌طور گسترده مورد استفاده قرار گرفته‌اند (وانگ، ۲۰۲۴؛ پریتی و همکاران، ۲۰۲۳؛ بریمن، ۲۰۰۱). با این حال، الگوریتم‌های درخت تصمیم به دلیل ساختار ساده، قابلیت تفسیر بالا و کارایی مناسب، همچنان جایگاه ویژه‌ای در طبقه‌بندی داده‌های سنجش‌ازدور دارند. الگوریتم *C5.0* به‌عنوان نسخه توسعه‌یافته *C4.5* (کوئینلن، ۱۹۹۳) با برخورداری از سرعت پردازش بالاتر، مصرف حافظه کمتر، قابلیت انتخاب ویژگی و هرس مبتنی بر کاهش خطا، توان بالقوه بالایی در طبقه‌بندی داده‌های پیچیده دارد (عبدالله و همکاران، ۲۰۲۱). افزون بر این، امکان ادغام این الگوریتم با تکنیک بوستینگ، دقت طبقه‌بندی کلاس‌های با هم‌پوشانی طیفی بالا را به‌طور قابل ملاحظه‌ای افزایش می‌دهد. شواهد پژوهشی نشان می‌دهد که *C5.0* در انتخاب ویژگی‌های مؤثر و کاهش افزونگی اطلاعات عملکرد موفقی دارد (وانگ، ۲۰۲۴). برای مثال، گیو و همکاران (۲۰۲۱) گزارش کردند که این الگوریتم در نقشه‌برداری آسیب‌پذیری زمین‌لغزش، دقتی معادل ۰.۸۸۳ کسب کرده که بالاتر از *SVM* و شبکه‌های بیزی بوده است. هرچند الگوریتم‌های ساده‌تری مانند *Tree* و *Rpart* در برخی کاربردها عملکرد قابل قبولی ارائه می‌دهند، اما توانایی *C5.0* در مدیریت داده‌های بزرگ و افزایش دقت طبقه‌بندی، آن را به گزینه‌ای مناسب‌تر برای مطالعات دقیق کاربری اراضی تبدیل کرده است.

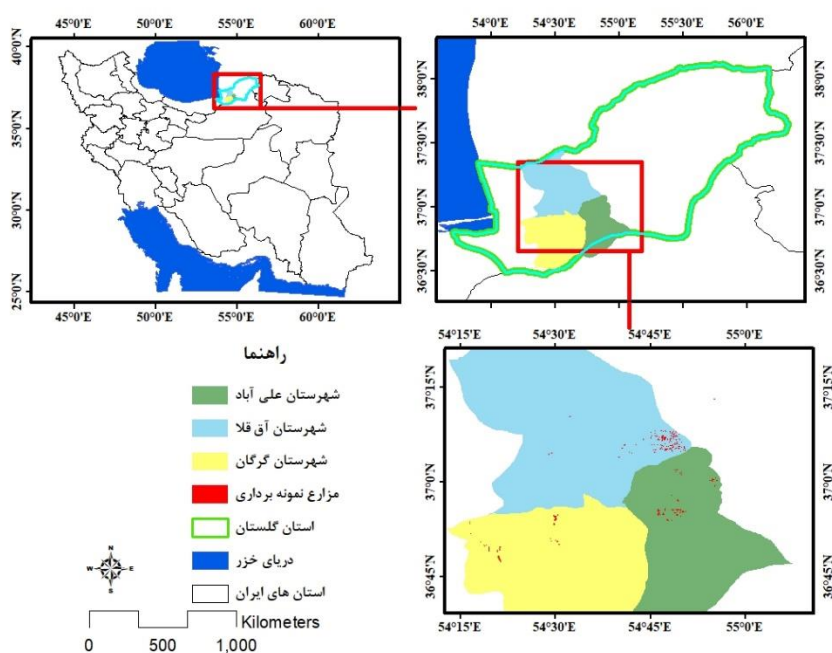
با توجه به پیچیدگی الگوهای طیفی و شباهت فنولوژیک محصولات مختلف، به‌ویژه در مقیاس مزرعه‌ای، استفاده از الگوریتم‌هایی که توانایی استخراج الگوهای معنادار از داده‌های چندزمانه را داشته باشند، امری ضروری است. بر این اساس، پژوهش حاضر با



شکل ۱- مراحل انجام پژوهش

سمنان و از غرب به استان مازندران و دریای خزر محدود می‌شود. از نظر توپوگرافی این استان از نقاط کوهستانی و جلگه‌ای تشکیل شده که قسمت جنوب و شرق را مناطق کوهستانی در بر گرفته که به طور موازی و با جهت تقریباً شرقی - غربی بوده و ادامه سلسله جبال البرز می‌باشد. بخش‌های مورد مطالعه از اقلیم‌های نیمه مرطوب، مدیترانه‌ای و نیمه خشک تشکیل شده‌اند. متوسط سالانه دمای هوا در این منطقه حدود ۱۷ درجه سانتی‌گراد متوسط بارش سالانه آنها بین ۶۰۰ تا ۳۰۰ میلیمتر است که بخش‌های شمالی آن از اقلیمی گرم و خشک‌تر برخوردار است.

**معرفی منطقه مورد مطالعه:** تحقیق حاضر در بخش‌هایی از مزارع استان گلستان واقع در شهرستان‌های گرگان، علی‌آباد و آق‌قلا اجرا می‌شود (شکل ۲). استان گلستان از نظر موقعیت جغرافیایی بین ۳۶ درجه و ۲۴ دقیقه تا ۳۸ درجه و ۵ دقیقه عرض شمالی و ۵۳ درجه و ۵۱ دقیقه تا ۵۶ درجه و ۱۴ دقیقه طول شرقی از نصف‌النهار گرینویچ قرار گرفته است. مساحت استان ۲۰۴۳۸ کیلومتر مربع بوده که معادل ۱/۳ درصد مساحت کل کشور را دارا می‌باشد. این استان از شمال به کشور ترکمنستان، از شرق به استان خراسان شمالی، از جنوب به استان



شکل ۲- موقعیت اراضی مورد مطالعه

### داده‌های مورد استفاده

در این پژوهش از دو دسته داده، یکی نمونه‌برداری‌های میدانی و دیگری تصاویر ماهواره‌ای استفاده شد. **داده‌های میدانی:** به منظور تهیه داده‌های مرجع قابل اعتماد برای آموزش و ارزیابی مدل‌های طبقه‌بندی، بازدیدهای میدانی گسترده‌ای در سطح منطقه مورد مطالعه در اواخر بهار و اوایل تابستان ۱۴۰۳ انجام شد. این بازدیدها با هدف شناسایی دقیق نوع کشت و ثبت اطلاعات مکانی مزارع محصولات زراعی اصلی شامل پنبه، برنج، سویا، توتون، آفتابگردان و ذرت صورت گرفت. نمونه‌های میدانی گردآوری شده (۵۷۳ مزرعه) به‌عنوان واقعیت زمینی نقش اساسی در افزایش دقت و اعتبار نتایج طبقه‌بندی مبتنی بر تصاویر ماهواره‌ای ایفا می‌کنند. در هر بازدید، موقعیت مکانی هر مزرعه با استفاده از دستگاه *GPS* دستی ثبت گردید و نقاط نمونه‌برداری به‌گونه‌ای انتخاب شدند که نماینده شرایط متنوع کشت، اندازه قطعات زراعی و پراکنش مکانی محصولات در منطقه باشند. در نهایت، داده‌های میدانی گردآوری شده به‌عنوان مبنای اصلی آموزش الگوریتم درخت تصمیم *C5.0* مورد استفاده قرار گرفتند. اتکای این پژوهش بر داده‌های میدانی دقیق و هم‌زمان با دوره رشد محصولات، موجب افزایش قابلیت اعتماد نتایج و کاهش عدم قطعیت در شناسایی و نقشه‌برداری محصولات زراعی در مقیاس مزرعه‌ای شده است.

**داده‌های ماهواره‌ای:** با توجه به اهداف تعریف‌شده در این پژوهش، از تصاویر ماهواره‌ای لندست ۸ و ۹ به‌عنوان منبع اصلی داده‌های سنجش‌ازدور استفاده شد. این ماهواره‌ها به‌سبب سنسورهای با قدرت تفکیک مکانی مناسب و پوشش زمانی منظم مجهز بوده و به‌طور گسترده برای پایش پوشش زمین، شناسایی محصولات کشاورزی و بررسی عوامل محیطی در مقیاس‌های منطقه‌ای مورد استفاده قرار می‌گیرند. تصاویر ماهواره‌ای لندست به دلیل دسترسی آزاد، پیوستگی زمانی بلندمدت و کیفیت طیفی مناسب، یکی از معتبرترین منابع داده‌ای در مطالعات کشاورزی و محیطی محسوب می‌شوند. تصاویر مورد استفاده از پایگاه داده سازمان زمین‌شناسی ایالات متحده آمریکا (*USGS*) و از طریق سامانه *Earth Explorer* دریافت شدند. در فرآیند جست‌وجو و دانلود داده‌ها، محدوده مورد مطالعه با استفاده از مختصات جغرافیایی، فایل برداری یا فرمت *Google Earth (KML)* تعریف شد. همچنین امکان تعیین موقعیت منطقه بر اساس سیستم مسیر-گذر ماهواره لندست فراهم است؛ به‌طوری که منطقه مورد مطالعه در این پژوهش در مسیر ۳۴ و گذر ۱۶۳ قرار

دارد. ماهواره‌های لندست ۸ و ۹ به دو سنجنده اصلی شامل سنجنده تصویربردار عملیاتی زمین ( $OLI^1$ ) و سنجنده فرورسرخ حرارتی ( $TIRS^2$ ) مجهز هستند. ترکیب داده‌های این دو سنجنده امکان دستیابی به اطلاعات طیفی و حرارتی را به‌طور هم‌زمان فراهم می‌سازد و نسبت به نسل‌های پیشین، به‌ویژه لندست ۷، بهبود قابل توجهی در کیفیت داده‌ها ایجاد کرده است. این ویژگی‌ها سبب شده است که تصاویر لندست ۸ و ۹ به‌طور گسترده در برآورد شاخص‌های پوشش گیاهی نظیر  $NDVI$ ، شاخص‌های تنش آبی، دمای سطح زمین و تخمین تبخیر-تعرق محصولات زراعی مورد استفاده قرار گیرند. مجموعه داده‌های عملیاتی زمین و حرارتی شامل ۱۱ باند طیفی است که ۷ باند آن با سنجنده  $ETM^+$  ماهواره لندست ۷ اشتراک طیفی دارند. پس از انتخاب مجموعه ماهواره‌ای مورد نظر و بازه زمانی مناسب، تصاویر به‌صورت فایل‌های با فرمت  $GeoTIFF$  دانلود و برای مراحل بعدی پردازش آماده شدند. در انتخاب تصاویر، حداکثر تلاش بر استفاده از صحنه‌هایی با کمترین درصد پوشش ابرناکی صورت گرفت تا اثر ابر و سایه آن بر دقت استخراج اطلاعات طیفی به حداقل برسد.

### شاخص‌های طیفی

استفاده از شاخص‌های طیفی امکان استخراج اطلاعات کمی دقیق‌تری از ویژگی‌های پوشش زمین و شرایط محیطی را فراهم می‌سازد و نقش مهمی در بهبود تحلیل‌های سنجش‌ازدور ایفا می‌کند (جنسن، ۲۰۱۶). مطابق با یافته‌های سانگ و همکاران (۲۰۰۱)، اگرچه اتمسفر باعث پراکندگی و جذب سیگنال‌های تابشی می‌شود، اما به‌کارگیری روش‌های تصحیح (مانند تفریق جسم تاریک) با کاهش تفاوت‌های بازتابی در طول زمان، دقت طبقه‌بندی و پایش تغییرات زمینی را به‌طور قابل توجهی افزایش می‌دهد. در این پژوهش، ماتریس داده‌های ورودی شامل مجموعه‌ای از شاخص‌های طیفی استخراج‌شده از تصاویر ماهواره‌ای لندست، در محیط نرم‌افزار  $ENVI$  (نسخه کلاسیک) تهیه شد (جدول ۱). افزون بر شاخص‌های طیفی، باندهای طیفی منفرد و نسبت‌های باندی منتخب نیز به‌عنوان سایر متغیرهای مستقل در فرآیند مدل‌سازی مورد استفاده قرار گرفتند تا توان تفکیک طیفی میان کلاس‌های کاربری اراضی افزایش یابد. استفاده از شاخص‌ها و نسبت‌های طیفی علاوه بر افزایش توان تفکیک کلاس‌های پوشش زمین، می‌تواند به‌طور مؤثری بخشی از اثرات ناخواسته اتمسفری، تغییرات روشنایی و تفاوت‌های شرایط تصویربرداری را کاهش دهد. این شاخص‌ها که معمولاً بر پایه نسبت یا ترکیب باندهای طیفی تعریف می‌شوند، حساسیت کمتری نسبت به تغییرات یکنواخت تابش، پراکندگی اتمسفری و اثرات زاویه خورشید سنجنده دارند و در نتیجه، پایداری بیشتری در تحلیل‌های چندزمانه فراهم می‌کنند.

**برازش داده‌ها و مدل‌سازی با استفاده از مدل درخت تصمیم C5:** مدل درخت تصمیم یکی از روش‌های پرکاربرد در مسائل طبقه‌بندی و پیش‌بینی کلاس متغیرها محسوب می‌شود. در این مدل، داده‌ها بر اساس متغیرهای توضیحی به مجموعه‌ای از گره‌ها تقسیم می‌شوند به‌گونه‌ای که هر گره دارای بیشترین همگنی درون‌کلاسی باشد (بريمن و همکاران، ۱۹۸۴). از جمله نخستین الگوریتم‌های درخت تصمیم، الگوریتم  $ID3$  است که در سال ۱۹۷۹ توسط کوئینلن معرفی شد. این الگوریتم با وجود سادگی، دارای محدودیت‌هایی از جمله اریب نسبت به متغیرهای با تعداد مقادیر زیاد و ناتوانی در مدیریت داده‌های گمشده بود. در راستای رفع این نواقص، نسخه بهبودیافته‌ای از این الگوریتم با عنوان  $C4.5$  در سال ۱۹۹۳ توسعه یافت که توانایی مدیریت داده‌های گمشده و کاهش اریب را دارا می‌باشد (کوئینلن، ۱۹۹۳).

<sup>1</sup> OLI: Operational Land Imager

<sup>2</sup> TIRS: Thermal Infrared Sensor

جدول ۱- شاخص‌های گیاهی منتخب و فرمول محاسبه در این تحقیق

شاخص طیفی	نام	رابطه	مرجع
NDVI	شاخص پوشش گیاهی نرمال شده تفاضلی	$(NIR - Red) / (NIR + Red)$	راس و همکاران، ۱۹۷۴
SAVI	شاخص پوشش گیاهی تعدیل شده با خاک	$1.5(NIR - Red) / (NIR + Red + 0.5)$	هیوت، ۱۹۸۸
RVI	نسبت کلروفیل شاخص پوشش گیاهی	$NIR/Red$	جردن، ۱۹۶۹
DVI	شاخص گیاهی تفاضلی	$NIR - Red$	استیون، ۱۹۹۸
IPVI	شاخص درصد پوشش گیاهی مادون قرمز	$NIR/(NIR+RDE)$	کریبن، ۱۹۹۰
EVI	شاخص گیاهی تعدیل شده	$2.5(NIR - Red) / (NIR + 6Red - 7.5Blue + 1)$	هیوت، ۲۰۰۲
LAI	شاخص سطح برگ	$3.618EVI - 0.118$	تیواری، ۲۰۰۳
GNDVI	شاخص پوشش گیاهی نرمال شده با سبزی برای تفاضل	$(NIR - Green) / (NIR + Green)$	هیوت، ۲۰۰۲
NGBDI	شاخص تفاضلی سبزی نرمال شده	$(GREEN-BLUE)/(GREEN+BLUE)$	هیوت، ۲۰۰۲
NDWI	شاخص تفاضل آب نرمال شده	$(NIR - SWIR1) / (NIR + SWIR1)$	گانو، ۱۹۹۶
CI	شاخص رنگ	$(NIR/GREEN) - 1$	اسمیت و همکاران، ۲۰۰۸
BI	شاخص روشنایی	$(GREEN2+RED2)0.5$	گانو، ۱۹۹۶
NMDI	شاخص اختلاف خشکسالی چند بانده نرمال شده	$(NIR - (SWIR1 - SWIR2)) / (NIR + (SWIR1 + SWIR2))$	وانگ و کیو، ۲۰۰۷
OSAVI	شاخص پوشش گیاهی اصلاح شده با خاک بهینه شده	$1.16(NIR - Red) / (NIR + Red + 0.16)$	روندآکس و همکاران، ۱۹۹۶

(۱)

$$Gain = Entropy(p) - \sum_{i=1}^k \frac{n_i}{n} Entropy(i)$$

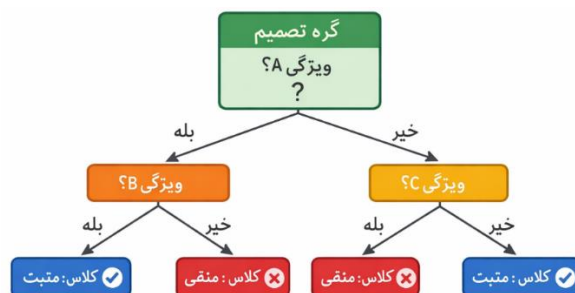
که در آن گره والد  $P$  به اجزاء  $K$  تقسیم می‌شود و  $n_i$  هم تعداد رکوردهای موجود در جزء  $i$  می‌باشد. به این ترتیب شباهت میان داده‌های موجود در یک گره و تفاوت میان مجموعه داده‌های موجود در گره‌ها بیشینه می‌شود. الگوریتم  $C5$  از روش هرس بدبینانه برای حذف طبقه‌بندی اشتباه استفاده می‌کند.

(۲)

$$Entropy(t) = - \sum_j p(j|t) \log_2 p(j|t)$$

$p(j|t)$  فراوانی نسبی کلاس  $j$  در گره  $t$  می‌باشد. عملیات مدل‌سازی در نرم‌افزار *Rapid Miner* صورت می‌گیرد (هان و همکاران، ۲۰۱۲؛ شانون، ۱۹۴۸).

در ادامه، الگوریتم پیشرفته‌تری با عنوان  $C5$  معرفی شد که بهبودهای قابل توجهی از نظر سرعت پردازش، دقت طبقه‌بندی و کارایی مدل نسبت به نسخه‌های پیشین ارائه می‌دهد. الگوریتم  $C5$  برای ارزیابی تقسیمات در هر گره، از معیار نسبت بهره اطلاعاتی (رابطه ۱) استفاده می‌کند که موجب انتخاب بهینه متغیرهای تفکیک‌کننده می‌شود (کوئینلن، ۱۹۹۶). در این پژوهش، به منظور استخراج ارتباط میان نوع محصول و پیش‌بینی الگوی پراکنش آن، از الگوریتم  $C5$  استفاده شده است. درخت تصمیم‌گیری در این الگوریتم، از میان متغیرهای محیطی، متغیری را انتخاب می‌کند که بیشترین بهره اطلاعاتی را برای تفکیک کلاس‌های محصول داشته باشد. همچنین، برای ارزیابی میزان سودمندی متغیر انتخاب شده، از معیار کاهش آنتروپی (رابطه ۲) استفاده می‌شود که بیانگر میزان کاهش عدم قطعیت در فرآیند طبقه‌بندی است.



شکل ۳- نمودار مدل C5

بر این اساس از معیارهایی همچون صحت یا پوشش نیز استفاده می‌کنند.

صحت یعنی از کل پیش‌بینی‌هایی که مدل طبقه‌بندی در مورد یک کلاس به‌خصوص انجام داده چند مورد درست بوده است.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

فراخوانی یعنی از کل نمونه‌هایی که از یک کلاس به‌خصوص در مجموعه داده‌ها موجود بوده، چند مورد توسط مدل طبقه‌بندی شناسایی شده است. تمرکز اصلی معیار فراخوانی بر خلاف معیار دقت بر روی داده‌هایی است که واقعاً «بلی» بوده‌اند.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$

### بحث و نتایج

در این بخش، نتایج حاصل از پردازش تصاویر ماهواره‌ای لندست، محاسبه شاخص‌های طیفی و اجرای مدل طبقه‌بندی مبتنی بر الگوریتم درختی C5.0 ارائه و تحلیل شده است. با توجه به هدف اصلی پژوهش که ارزیابی کارایی شاخص‌های طیفی منتخب در تفکیک الگوهای کشت و سنجش توان مدل در شناسایی کلاس‌های مختلف کاربری اراضی است، یافته‌ها ابتدا در سه تاریخ ۶ آگوست ۲۰۲۴، ۳۰ آگوست ۲۰۲۴ و ۲۳ سپتامبر ۲۰۲۴ که به ترتیب معادل با تاریخ‌های شمسی ۱۴۰۳/۰۵/۱۶، ۱۴۰۳/۰۶/۰۹ و ۱۴۰۳/۰۷/۰۲ بوده؛ سازمان‌دهی

**معیارهای ارزیابی خطا:** در وظایف طبقه‌بندی در یادگیری ماشین، یک ماتریس درهم‌ریختگی<sup>۱</sup> به عنوان جدولی برای خلاصه کردن عملکرد مدل عمل می‌کند که شامل تعداد پیش‌بینی‌های درست مثبت ( $TP^2$ )، درست منفی ( $TN^3$ )، مثبت کاذب ( $FP^4$ ) و منفی کاذب ( $FN^5$ ) است. برای طبقه‌بندی چندکلاس، این ماتریس به شکل مربعی با ابعادی برابر با تعداد کلاس‌ها می‌شود که هر سلول نمایانگر تعداد نمونه‌ها برای ترکیب برچسب‌ها و کلاس‌های پیش‌بینی‌شده است. ردیف‌ها نشان‌دهنده کلاس‌های واقعی هستند، در حالی که ستون‌ها نشان‌دهنده کلاس‌های پیش‌بینی‌شده می‌باشند. تحلیل این ماتریس دیدگاهی نسبت به نقاط قوت و ضعف مدل در تمایز بین کلاس‌های خاص ارائه می‌دهد. بر اساس نتایج این ماتریس می‌توان از معیارهای کمی شامل درستی<sup>۶</sup>، دقت<sup>۷</sup> و فراخوانی<sup>۸</sup> نیز برای ارزیابی مدل طبقه‌بندی نیز استفاده کرد. درستی یعنی از کل پیش‌بینی‌های انجام شده توسط مدل طبقه‌بندی کننده چند مورد درست بوده است. معیار درستی در نمونه‌هایی که توزیع یک کلاس در آن‌ها بسیار زیادتر از کلاس‌های دیگر باشد، معیار مناسبی نیست.

(۳)

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN}$$

<sup>5</sup> FN: False Negative

<sup>6</sup> Accuracy

<sup>7</sup> Precision

<sup>8</sup> Recall

<sup>1</sup> Confusion matrix

<sup>2</sup> TP: True Positive

<sup>3</sup> TN: True Negative

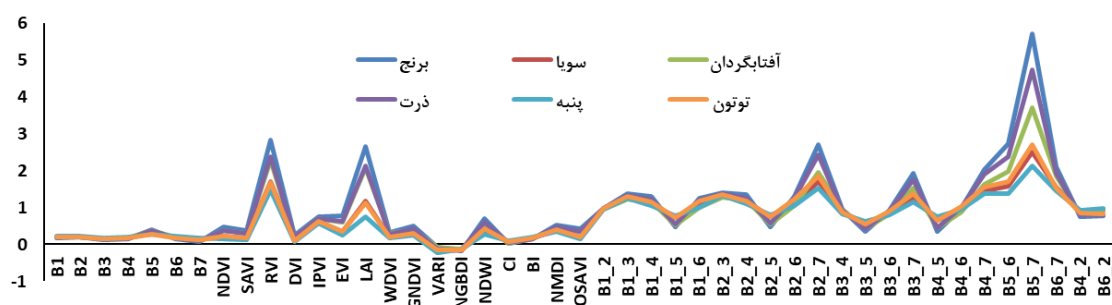
<sup>4</sup> FP: False Positive

و فعالیت فتوسنتزی قوی‌تر آن‌ها در مقطع زمانی مطالعه است. در مقابل، سویا و ذرت رفتار طیفی نزدیک‌تری به یکدیگر دارند که می‌تواند توجیه‌کننده میزان اختلاط بالاتر این دو کلاس در ماتریس درهم‌ریختگی باشد. شاخص‌های حساس به رطوبت و ساختار پوشش نظیر  $NDWI$ ،  $MSI$  و  $CI$  نیز تفاوت نسبی میان محصولات را برجسته می‌کنند، به‌ویژه در تفکیک برنج به‌دلیل شرایط رطوبتی خاص آن. همچنین، باندهای بازنمونه‌برداری شده ( $B1-2$  تا  $B6$ ) (زروندی مشابه اما با دامنه نوسان متفاوت را نشان می‌دهند که حاکی از پایداری الگوی طیفی محصولات در طول باندها و نقش مؤثر ترکیب ویژگی‌های چندباندی در فرآیند تصمیم‌گیری مدل  $C5$  است. در مجموع، این نمودار نشان می‌دهد که شاخص‌های پوشش گیاهی و باندهای نزدیک‌فروسرخ بیشترین سهم را در تفکیک محصولات زراعی داشته‌اند و مبنای اصلی گره‌های تصمیم در مدل درختی  $C5$  را تشکیل داده‌اند، در حالی که هم‌پوشانی طیفی برخی محصولات، به‌ویژه سویا و ذرت، همچنان یک چالش اصلی در طبقه‌بندی مبتنی بر داده‌های لندست محسوب می‌شود.

شده‌اند. عملکرد مدل  $C5.0$  تشریح شده و ساختار درخت تصمیم و قواعد استخراج‌شده از آن مورد تحلیل قرار گرفته است. در نهایت، نتایج در این ۳ تاریخ مقایسه و نتیجه‌گیری نهایی انجام شد.

### تفسیر نتایج طبقه‌بندی با مدل $C5$

**تفسیر نتایج طبقه‌بندی با مدل  $C5$  در تاریخ ۱۶ مرداد ۱۴۰۳:** نمودار ارائه‌شده الگوی تغییرات بازتاب طیفی و شاخص‌های استخراج‌شده از تصاویر لندست را برای شش محصول زراعی شامل برنج، سویا، ذرت، آفتابگردان، توتون و پنبه نشان می‌دهد (شکل ۴) و مبنای مناسبی برای تبیین منطق تفکیک در مدل درختی  $C5.0$  فراهم می‌سازد. در باندهای مرئی- $(B1-3)$  اختلاف معنی‌داری میان محصولات مشاهده نمی‌شود که بیانگر محدودیت این ناحیه طیفی در تفکیک محصولات زراعی است. با این حال، از باند نزدیک‌فروسرخ و شاخص‌های مبتنی بر آن، به‌ویژه  $NDVI$ ،  $RVI$ ،  $DVI$ ،  $EVI$  و  $LAI$ ، الگوهای تفکیک‌پذیرتری آشکار می‌شود؛ به‌طوری‌که برنج و پنبه بیشترین مقادیر را در این شاخص‌ها نشان می‌دهند که منعکس‌کننده تراکم پوشش گیاهی بالاتر



شکل ۴- منحنی بازتاب طیفی، شاخص‌ها و نسبت‌های طیفی محصولات مختلف مستخرج از تصاویر لندست در تاریخ ۱۶/۰۵/۱۴۰۳

حاکی از تفکیک مناسب الگوهای طیفی و فنولوژیکی این محصولات در داده‌های لندست است. در مقابل، کلاس سویا با صحت ۶۶/۸۲ درصد و فراخوانی ۵۲/۴۶ درصد و همچنین ذرت با فراخوانی ۶۳/۰۵ درصد، بیشترین میزان اختلاط را با سایر کلاس‌ها نشان می‌دهند؛ که می‌تواند ناشی از شباهت‌های طیفی

نتایج حاصل از اجرای مدل درختی  $C5$  نشان می‌دهد (جدول ۲) تشخیص محصولات غالب، به‌ویژه پنبه و شالی، از توان تفکیک بالاتری برخوردار است؛ به‌طوری‌که صحت طبقه‌بندی این دو کلاس به ترتیب ۸۸/۱۲ درصد و ۸۵/۳۴ درصد و میزان فراخوانی آن‌ها ۸۹/۶۷ درصد و ۹۰/۱۸ درصد به‌دست آمده است که

شاخص‌ها و باندهای طیفی لندست، قابلیت مناسبی در شناسایی محصولات زراعی اصلی دارد، اما بهبود تفکیک کلاس‌های با رفتار طیفی مشابه مستلزم استفاده از داده‌های چندزمانه متراکم‌تر، شاخص‌های فنولوژیکی و یا تلفیق داده‌های سنجنده‌های با تفکیک طیفی بالاتر است.

هم‌پوشانی زمانی مراحل رشد و محدودیت تفکیک طیفی سنجنده لندست در برخی باندها باشد. کلاس آفتابگردان نیز با وجود تعداد نمونه کمتر، دقت متوسطی برابر ۸۲/۳۵ درصد و Recall برابر ۸۰/۹۲ درصد را ثبت کرده است. به‌طور کلی، این ماتریس درهم‌ریختگی نشان می‌دهد که مدل C5 در ترکیب با

جدول ۲- نتایج تفکیک محصولات زراعی با استفاده از مدل C5 در تاریخ ۱۴۰۳/۰۵/۱۶

		واقعی							
		دقت کلی:	پنبه	توتون	آفتابگردان	ذرت	سویا	شالی	درصد صحت کلاس
پیش‌بینی شده	پنبه	۸۳/۶۵ درصد	۱۲۶۸	۶۱	۱۵	۲۳	۵۲	۱۰	۸۸/۱۲
	توتون		۶۷	۷۳۳	۷	۲۱	۶۱	۱۷	۸۰/۹۱
	آفتابگردان		۷	۵	۱۴۰	۳	۱۵	۰	۸۲/۳۵
	ذرت		۱۳	۳	۰	۱۸۶	۶	۲۸	۷۸/۸۱
	سویا		۳۵	۱۸	۳	۹	۱۴۹	۹	۶۶/۸۲
	شالی		۲۴	۲۵	۸	۴۳	۱	۵۸۸	۸۵/۳۴
	درصد پوشش (بازیابی) کلاس		۸۹/۶۷	۸۶/۷۵	۸۰/۹۲	۶۳/۰۵	۵۲/۴۶	۹۰/۱۸	

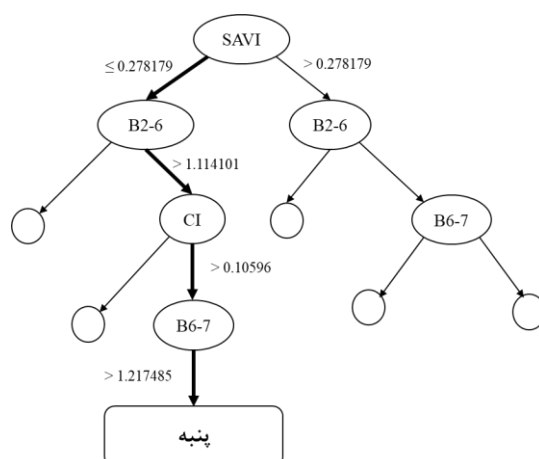
عمدتاً با مقادیر پایین CI و نسبت‌های خاص SWIR (نشان‌دهنده رطوبت بالاتر گیاه) شناسایی شده، در حالی که ذرت با مقادیر بالاتر SAVI و نسبت‌های متمایز NIR-SWIR از سایر محصولات تفکیک می‌شود. ساختار سلسله‌مراتبی این درخت نشان می‌دهد که مدل C5 با بهره‌گیری هم‌زمان از شاخص‌های پوشش گیاهی، شاخص‌های کلروفیل و نسبت‌های باندهای مرتبط با رطوبت، قادر است تفاوت‌های فنولوژیک و فیزیولوژیک محصولات زراعی را به‌صورت تفسیرپذیر مدل‌سازی کرده و تفکیک قابل‌اعتمادی بین محصولات استراتژیک منطقه فراهم آورد.

تحلیل هم‌زمان منحنی‌های بازتاب طیفی محصولات زراعی و ساختار مدل درخت تصمیم C5 استخراج‌شده از داده‌های لندست نشان می‌دهد که تفکیک موفق محصولات عمدتاً مبتنی بر تفاوت‌های فیزیولوژیک، فنولوژیک و رطوبتی پوشش گیاهی است. منحنی‌های طیفی بیانگر بازتاب بالاتر محصولات متراکم نظیر برنج و ذرت در بلند مادون‌قرمز نزدیک و

نمودار تقلیل یافته درختی نیز نشان می‌دهد که شاخص SAVI به‌عنوان مهم‌ترین متغیر تفکیک‌کننده در گره ریشه انتخاب شده است که بیانگر نقش کلیدی تراکم پوشش گیاهی و کاهش اثر پس‌زمینه خاک در شناسایی محصولات زراعی است. مقادیر پایین ( $SAVI \leq 0/278$ ) عمدتاً به محصولات با پوشش گیاهی ضعیف‌تر یا مراحل ابتدایی رشد اختصاص یافته و در این شاخه، نسبت بلندی B2/B6 آبی به SWIR به‌عنوان شاخص حساس به ویژگی‌های سطحی و رطوبت گیاه-خاک نقش تعیین‌کننده‌ای در تفکیک پنبه و آفتابگردان ایفا می‌کند. در ادامه، شاخص CI نسبت B6/B7 به‌طور مؤثری برای تمایز محصولات با تفاوت در محتوای کلروفیل و تنش آبی، به‌ویژه در جداسازی پنبه از آفتابگردان، به‌کار گرفته شده‌اند. در شاخه مربوط به مقادیر بالاتر ( $SAVI > 0/278$ )، که معرف پوشش‌های گیاهی متراکم‌تر است، ترکیب نسبت‌های B2/B6 و B6/B7 به‌همراه CI و آستانه‌های ثانویه SAVI، امکان تفکیک دقیق محصولات برنج، ذرت و سویا را فراهم کرده است؛ به‌طوری‌که برنج

رطوبت بالا، را فراهم کرده است. همچنین، به‌کارگیری شاخص  $CI$  در گره‌های میانی درخت با اختلافات مشاهده‌شده در باندهای مرئی و شاخص‌های کلروفیل هم‌خوانی داشته و به تفکیک محصولات با سطح کلروفیل و مراحل رشد متفاوت، نظیر ذرت و سویا، کمک کرده است. این هم‌خوانی ساختاری میان الگوهای بازتاب طیفی و قوانین تصمیم مدل  $C5$  نشان می‌دهد که ترکیب باندهای  $NIR$  و  $SWIR$  به‌همراه شاخص‌های پوشش گیاهی و نسبت‌های بلندی، چارچوبی تفسیرپذیر و کارآمد برای طبقه‌بندی محصولات زراعی در تصاویر ماهواره‌ای لندست فراهم می‌آورد. در شکل (۵) نمودار درختی تقلیل یافته برای شناسایی مزارع پنبه در تاریخ ۱۶ مرداد ۱۴۰۳ ارائه شده است.

افت محسوس آن‌ها در باندهای  $SWIR$  به‌دلیل محتوای رطوبت بالاتر گیاه هستند، در حالی که پنبه و آفتابگردان بازتاب کمتر در  $NIR$  و مقادیر بالاتری در  $SWIR$  از خود نشان می‌دهند که نشان‌دهنده ساختار تاج بازتر و تحمل بیشتر به تنش رطوبتی است. این الگوهای طیفی به‌طور مستقیم در ساختار درخت تصمیم بازتاب یافته‌اند، به‌گونه‌ای که شاخص  $SAVI$  به‌عنوان گره ریشه، نقش کلیدی تراکم پوشش گیاهی و کاهش اثر خاک را در جداسازی اولیه محصولات آشکار می‌سازد. در ادامه، استفاده از نسبت باندی  $B2/B6$  و نسبت  $B6/B7$  که در نمودار طیفی نیز بیشترین اختلاف بین محصولات را نشان می‌دهند، امکان تفکیک محصولات با رژیم‌های آبی متفاوت، به‌ویژه شناسایی دقیق برنج به‌عنوان محصول وابسته به



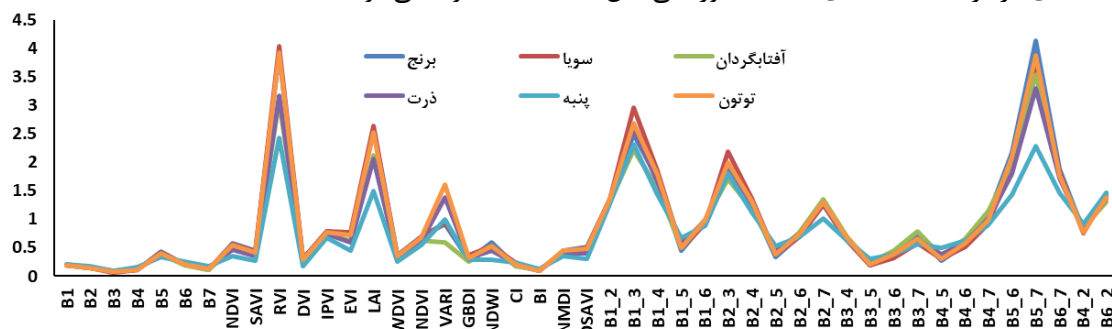
شکل ۵- نمودار درختی تقلیل یافته برای شناسایی مزارع پنبه در تاریخ ۱۶/۰۵/۱۴۰۳

$SAVI$ ،  $NDVI$ ،  $DVI$ ،  $EVI$  و  $LAI$ ، بیشترین قدرت تمایز را نشان می‌دهند؛ به‌طوری‌که برنج و سویا مقادیر بالاتری را در این شاخص‌ها ثبت کرده‌اند که بازتاب‌دهنده تراکم پوشش گیاهی، سطح برگ بالاتر و فعالیت فتوسنتزی قوی‌تر در مقطع زمانی مورد مطالعه است. ذرت و آفتابگردان رفتار طیفی نسبتاً مشابهی، به‌ویژه در شاخص‌های گیاهی، از خود نشان می‌دهند که می‌تواند بیانگر شباهت ساختار تاج پوشش و مراحل فنولوژیکی این محصولات باشد. شاخص‌های حساس به رطوبت و تنش گیاهی نظیر  $NDWI$ ،  $MSI$  و  $CI$  نیز اختلاف مشخصی میان محصولات ایجاد کرده‌اند و نقش

تفسیر نتایج طبقه‌بندی با مدل  $C5$  در تاریخ ۹ شهریور ۱۴۰۳: الگوی پاسخ طیفی و شاخص‌های استخراج‌شده از تصاویر ماهواره‌ای لندست برای شش محصول زراعی برنج، سویا، ذرت، آفتابگردان، توتون و پنبه در تاریخ ۹ شهریور ماه بیانگر تفاوت‌های معنادار در رفتار طیفی و بیوفیزیکی این محصولات است (شکل ۶). در باندهای مرئی ( $B1-B3$ ) دامنه تغییرات میان کلاس‌ها محدود بوده و هم‌پوشانی قابل توجهی مشاهده می‌شود که نشان‌دهنده نقش ضعیف این باندها در تفکیک محصولات است. در مقابل، شاخص‌های مبتنی بر ناحیه نزدیک‌فروسرخ و ترکیبات طیفی، به‌ویژه

نمودار تأیید می‌کند که شاخص‌های پوشش گیاهی و باندهای نزدیک‌فروسرخ بیشترین سهم را در تفکیک محصولات زراعی داشته و مبنای اصلی تصمیم‌گیری در مدل درختی C5 را تشکیل داده‌اند، در حالی که هم‌پوشانی طیفی برخی محصولات همچنان چالشی اساسی در طبقه‌بندی مبتنی بر داده‌های لندست محسوب می‌شود.

مهمی در تفکیک محصولاتی با نیاز آبی متفاوت، به‌ویژه برنج، ایفا می‌کنند. همچنین، باندهای مشتق‌شده و بازنمونه‌برداری‌شده (BI-2) تا (B6-2) الگوی نوسانی مشابه اما با دامنه‌های متفاوت را نشان می‌دهند که بیانگر پایداری نسبی پاسخ طیفی محصولات در طول باندها و اهمیت استفاده از مجموعه‌ای از ویژگی‌های چندباندی در فرآیند طبقه‌بندی است. به‌طور کلی، این



شکل ۶- منحنی بازتاب طیفی، شاخص‌ها و نسبت‌های طیفی محصولات مختلف مستخرج از تصاویر لندست در تاریخ ۱۴۰۳/۰۶/۰۹

کمتر است. در مقابل، کلاس سویا ضعیف‌ترین عملکرد را ثبت کرده است؛ به‌طوری‌که کمترین مقدار بازیابی ۵۸/۸۰ درصد و صحت ۷۰/۱۷ درصد را دارد که نشان‌دهنده هم‌پوشانی طیفی و شباهت فنولوژیکی آن با ذرت و توتون است. کلاس ذرت نیز با صحت ۷۷/۷۰ درصد و بازیابی ۷۵/۵۹ درصد عملکردی متوسط داشته و بخشی از نمونه‌های آن با پنبه و سویا دچار اختلاط شده‌اند. به‌طور کلی، این ماتریس درهم‌ریختگی نشان می‌دهد که مدل در تشخیص محصولات با رفتار طیفی و شرایط رشد متمایز، به‌ویژه برنج و پنبه، موفق‌تر عمل کرده است، در حالی که تفکیک محصولاتی با شباهت طیفی بالا، به‌خصوص سویا، همچنان چالش‌برانگیز باقی مانده و نیازمند استفاده از داده‌های چندزمانه متراکم‌تر، شاخص‌های فنولوژیکی و یا تلفیق داده‌های سنجنده‌ای با تفکیک طیفی بالاتر است.

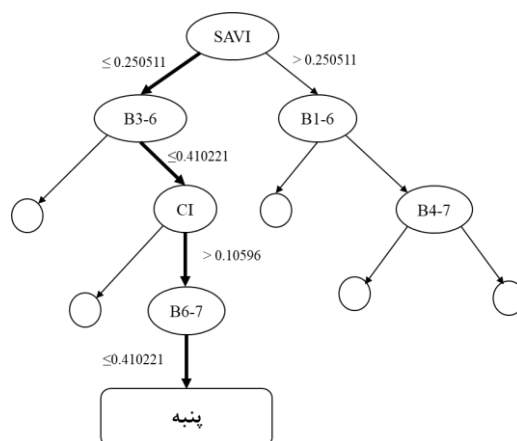
ماتریس درهم‌ریختگی ارائه‌شده (جدول ۳) عملکرد مدل درختی C5 را در تفکیک شش محصول زراعی با استفاده از تصاویر ماهواره‌ای لندست نشان می‌دهد که منجر به دقت کلی ۷۹/۸۸ درصد شده است و بیانگر کارایی متوسط تا قابل‌قبول مدل در شرایط مورد مطالعه است. نتایج نشان می‌دهد که کلاس برنج با بازیابی بالا ۸۷/۵۸ درصد و صحت مناسب ۷۹/۵۳ درصد به‌خوبی شناسایی شده است که می‌تواند ناشی از ویژگی‌های طیفی متمایز و شرایط رطوبتی خاص این محصول باشد. پنبه نیز با بازیابی ۸۶/۳۵ درصد عملکرد مناسبی در شناسایی نمونه‌های واقعی خود داشته است، هرچند صحت آن ۸۱/۶۷ درصد نشان‌دهنده مقداری اختلاط با سایر کلاس‌ها، به‌ویژه توتون و ذرت، است. کلاس آفتابگردان با صحت ۸۲/۶۳ درصد و بازیابی ۷۹/۷۷ درصد دقت نسبتاً پایداری را نشان می‌دهد که بیانگر تفکیک قابل‌قبول این محصول علی‌رغم تعداد نمونه

جدول ۳- نتایج تفکیک محصولات زراعی با استفاده از مدل C5 در تاریخ ۱۴۰۳/۰۶/۰۹

دقت کلی: ۷۹/۸۸ درصد	واقعی						درصد صحت کلاس
	پنبه	توتون	آفتابگردان	ذرت	سویا	شالی	
پنبه	۱۲۲۱	۱۲۷	۵	۳۸	۷۳	۳۱	۸۱/۶۷
توتون	۶۷	۶۰۶	۱۱	۱۷	۲۸	۲۹	۷۹/۹۵
آفتابگردان	۵	۱۹	۱۳۸	۱	۰	۴	۸۲/۶۳
ذرت	۴۰	۴	۰	۲۲۳	۱۵	۵	۷۷/۷۰
سویا	۱۶	۳۶	۱	۶	۱۶۷	۱۲	۷۰/۱۷
شالی	۶۵	۵۳	۱۸	۱۰	۱	۵۷۱	۷۹/۵۳
درصد پوشش (بازیابی) کلاس	۸۶/۳۵	۷۱/۷۲	۷۹/۷۷	۷۵/۵۹	۵۸/۸۰	۸۷/۵۸	

به‌همراه آستانه‌های ثانویه SAVI، اختلافات فنولوژیک و فیزیولوژیک بین محصولات را به‌خوبی منعکس می‌کند؛ به‌طوری‌که برنج عمدتاً در گره‌هایی با مقادیر پایین‌تر نسبت‌های SWIR و حساسیت بالاتر به رطوبت شناسایی شده، در حالی که ذرت با مقادیر بالاتر SAVI و نسبت‌های متمایز مرئی SWIR از سویا و آفتابگردان تفکیک شده است. حضور مکرر نسبت‌های بانندی مرتبط با SWIR در گره‌های میانی و پایانی درخت (شکل ۷) نشان می‌دهد که محتوای رطوبت گیاه و خاک عامل کلیدی در تفکیک محصولات زراعی منطقه بوده و ساختار این درخت تصمیم به‌خوبی هم‌راستا با الگوهای بازتاب طیفی استخراج‌شده از تصاویر لندست است، به‌گونه‌ای که مدل C5 ضمن حفظ تفسیرپذیری، قابلیت بالایی در شناسایی و جداسازی محصولات استراتژیک کشاورزی از خود نشان می‌دهد.

مدل درخت تصمیم C5 حاصل از ترکیب شاخص‌های پوشش گیاهی و نسبت‌های بانندی لندست نشان می‌دهد که شاخص SAVI بار دیگر به‌عنوان مهم‌ترین متغیر تفکیک‌کننده در گره ریشه انتخاب شده است که بیانگر نقش تعیین‌کننده تراکم پوشش گیاهی و تعدیل اثر پس‌زمینه خاک در مراحل اولیه طبقه‌بندی محصولات زراعی است. مقادیر پایین (۰/۲۵)  $SAVI \leq$  عمدتاً به محصولات با پوشش گیاهی ضعیف‌تر یا مراحل ابتدایی رشد اختصاص یافته و در این شاخه، نسبت بانندی B3/B6 سبز به SWIR1 و سپس B1/B2 ساحلی به آبی به‌عنوان شاخص‌هایی حساس به شرایط سطحی و رطوبت خاک، امکان تفکیک مؤثر پنبه از سایر محصولات را فراهم کرده‌اند. در شاخه مقادیر بالاتر ( $SAVI > 0.25$ )، که معرف پوشش‌های گیاهی متراکم‌تر است، ترکیب سلسله‌مراتبی نسبت‌های B3/B6، B2/B6 و B4/B7

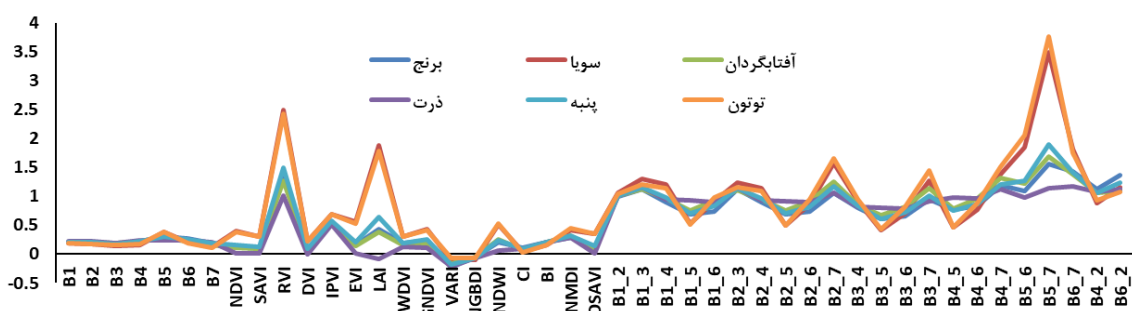


شکل ۷- نمودار درختی تقلیل یافته برای شناسایی مزارع پنبه در تاریخ ۱۴۰۳/۰۶/۰۹

کلاس‌ها را نشان می‌دهند؛ به‌گونه‌ای که توتون در این شاخص‌ها مقادیر بالاتری را ثبت کرده و بیانگر تراکم زیست‌توده و سطح برگ بیشتر در زمان برداشت تصویر است. برنج و پنبه نیز رفتار طیفی نسبتاً نزدیک اما با مقادیر بالاتر نسبت به سویا و ذرت دارند که این امر می‌تواند ناشی از تفاوت در ساختار تاج پوشش و وضعیت رطوبتی مزارع باشد. شاخص‌های حساس به رطوبت و تنش نظیر *NDWI*، *NMDI* و *MSI* نیز اختلاف نسبی میان محصولات را برجسته می‌کنند و به‌ویژه در تمایز برنج از سایر محصولات نقش مؤثری ایفا می‌نمایند. همچنین، باندهای بازمنه‌برداری شده (*BI-2* تا *B6*) (2) روندی یکنواخت اما با دامنه نوسان متفاوت را نشان می‌دهند که حاکی از پایداری الگوی پاسخ طیفی محصولات و اهمیت استفاده هم‌زمان از چندین باند و شاخص در فرآیند طبقه‌بندی است. در مجموع، این نمودار نشان می‌دهد که شاخص‌های گیاهی و باندهای نزدیک‌فروسرخ بیشترین سهم را در تفکیک محصولات زراعی داشته و مبنای اصلی تصمیم‌گیری در مدل درختی C5 را تشکیل داده‌اند. در حالی که هم‌پوشانی طیفی برخی محصولات، به‌ویژه سویا و ذرت، همچنان یکی از عوامل اصلی خطای طبقه‌بندی در داده‌های لندست محسوب می‌شود.

مدل یک مسیر کارآمد برای شناسایی پنبه با تراکم پوشش گیاهی کم تا متوسط پیدا کرده است. این نشان می‌دهد که ترکیب خاصی از بازتاب در باندهای ۳ و ۶ در این تاریخ یک نشانگر قوی برای پنبه است. برای پیکسل‌های با *SAVI* بالا، درخت پیچیده‌تر می‌شود و از ترکیبی از نسبت‌های باندی مختلف برای تفکیک محصولات استفاده می‌کند. کاهش جزئی در دقت کلی مدل در این تاریخ ممکن است به دلیل هم‌پوشانی طیفی بیشتر محصولات در این مرحله رشد باشد. با این حال صحت تشخیص پنبه در این تاریخ همچنان بالاست.

**تفسیر نتایج طبقه‌بندی با مدل C5 در تاریخ دوم مهر ۱۴۰۳:** نمودار حاضر (شکل ۸) الگوی تغییرات بازتاب طیفی باندها و شاخص‌های گیاهی استخراج‌شده از تصاویر ماهواره‌ای لندست را برای شش محصول زراعی برنج، سویا، ذرت، آفتابگردان، توتون و پنبه نشان می‌دهد و تفاوت‌های رفتاری این محصولات را در فضای ویژگی‌ها به‌خوبی آشکار می‌سازد. در باندهای مرئی (*BI-B3*) اختلاف محسوسی میان محصولات مشاهده نمی‌شود که بیانگر قدرت تفکیک محدود این ناحیه طیفی است. در مقابل، شاخص‌های مبتنی بر ترکیب باند قرمز و نزدیک‌فروسرخ، به‌ویژه *NDVI*، *RVI*، *DVI*، *EVI* و *LAI*، بیشترین دامنه تغییرات و تمایز بین



شکل ۸- منحنی بازتاب طیفی، شاخص‌ها و نسبت‌های طیفی محصولات مختلف مستخرج از تصاویر لندست در تاریخ ۱۴۰۳/۰۷/۰۲

ماهواره‌ای لندست است؛ به‌طوری‌که دقت کلی ۸۸/۱۰ درصد بیانگر بهبود قابل توجه کارایی مدل نسبت به سناریوهای قبلی است. نتایج حاکی از آن است که

ماتریس درهم‌ریختگی ارائه‌شده مطابق جدول (۴) نشان‌دهنده عملکرد بسیار مطلوب مدل درختی C5 در تفکیک شش محصول زراعی با استفاده از تصاویر

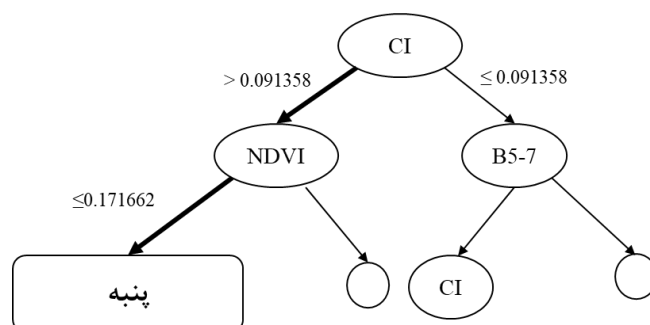
مقابل، کلاس سویا با وجود صحت قابل قبول ۸۶/۸۴ درصد، کمترین میزان بازیابی ۶۹/۷۲ درصد را دارد؛ که نشان‌دهنده باقی‌ماندن چالش هم‌پوشانی طیفی و فنولوژیکی آن با سایر محصولات، به‌ویژه ذرت و پنبه، است. به‌طور کلی، این ماتریس درهم‌ریختگی تأیید می‌کند که مدل C5 در ترکیب با باندها و شاخص‌های طیفی استخراج‌شده از تصاویر لندست، توان بالایی در شناسایی محصولات زراعی دارد و به‌ویژه در تفکیک کلاس‌های ذرت و آفتابگردان عملکرد بسیار مطلوبی ارائه می‌دهد؛ با این حال، بهبود تشخیص سویا همچنان نیازمند بهره‌گیری از داده‌های چندزمانه متراکم‌تر، شاخص‌های فنولوژیکی و یا تلفیق داده‌های سنجنده‌های با تفکیک طیفی بالاتر است.

کلاس‌های ذرت و آفتابگردان با صحت‌های بسیار بالا به‌ترتیب ۹۳/۷۳ درصد و ۹۲/۵۷ درصد و مقادیر بازیابی چشمگیر ۹۱/۱۹ درصد و ۹۳/۶۴ درصد به‌طور دقیق شناسایی شده‌اند که نشان‌دهنده تمایز مناسب ویژگی‌های طیفی و ساختاری این محصولات در داده‌های لندست است. کلاس برنج نیز با صحت ۸۸/۴۰ درصد و بازیابی ۹۰/۰۳ درصد عملکرد پایداری را ثبت کرده که می‌تواند ناشی از شرایط رطوبتی خاص و رفتار طیفی متمایز این محصول باشد. کلاس پنبه و توتون هر دو صحت یکسان ۸۷/۶۲ درصد و بازیابی‌های بالا به‌ترتیب ۹۰/۱۷ درصد و ۸۷/۱۰ درصد را نشان می‌دهند که بیانگر توان مناسب مدل در تفکیک این دو محصول با وجود شباهت نسبی طیفی است. در

جدول ۴- نتایج تفکیک محصولات زراعی با استفاده از مدل C5 در تاریخ ۱۴۰۳/۰۷/۰۲

درصد صحت کلاس	واقعی						
	شالی	سویا	ذرت	آفتابگردان	توتون	پنبه	دقت کلی: ۸۸/۱ درصد
۸۷/۶۲	۵۰	۴۱	۱۳	۵	۸۵	۱۲۷۵	پنبه
۸۷/۶۲	۱۴	۴۱	۶	۶	۷۳۶	۳۷	توتون
۹۲/۵۷	۰	۳	۴	۱۶۲	۳	۳	آفتابگردان
۹۳/۷۳	۰	۰	۲۶۹	۰	۴	۱۴	ذرت
۸۶/۸۴	۱	۱۹۸	۳	۰	۱۵	۱۱	سویا
۸۸/۴۰	۵۷۸	۱	۰	۰	۲	۷۴	شالی
	۹۰/۰۳	۶۹/۷۲	۹۱/۱۹	۹۳/۶۴	۸۷/۱۰	۹۰/۱۷	درصد پوشش (بازیابی) کلاس

پیش‌بینی شده



شکل ۹- نمودار درختی تقلیل یافته برای شناسایی مزارع پنبه در تاریخ ۱۴۰۳/۰۷/۰۲

خود را تغییر داده و شاخص رنگ را به عنوان مهم‌ترین متغیر برای تفکیک اولیه انتخاب می‌کند.

در اواخر فصل رشد، با بازشدن غوزه‌های پنبه و تغییر رنگ سایر محصولات، مدل به طور کامل منطبق

عمده‌ای از کلاس‌های زراعی را با صحت مناسب شناسایی کند. این الگو با گزارش‌های منتشرشده در ادبیات جهانی درباره کارایی درخت تصمیم در مطالعات پایش محصولات زراعی همسو است. برای مثال، ژانگ و همکاران (۲۰۲۵) و ژائو و همکاران (۲۰۱۱) در مطالعه‌ای بر روی اراضی کشاورزی پکن نشان دادند که الگوریتم C5 تونایی بالایی در تفکیک محصولات با رفتار فنولوژیک مشخص دارد و به‌ویژه در دوره رشد کامل گیاهان، از بسیاری از روش‌های کلاسیک نظیر *Maximum Likelihood* عملکرد بهتری ارائه می‌دهد. این نتیجه با یافته‌های تاریخ ۱۶ مرداد در پژوهش حاضر همخوان است، جایی که محصولات با بیشترین رشد رویشی، به‌خصوص پنبه و برنج، بالاترین صحت تفکیک را ثبت کرده‌اند.

علاوه بر این، نتایج به‌دست‌آمده برای تاریخ ۹ شهریور که مرحله گذار از رشد کامل به فاز رسیدگی را نشان می‌دهد، با مطالعاتی نظیر فریز و همکاران (۲۰۱۷) و پریتی و همکاران (۲۰۲۳) هم‌راستا است. این محققان گزارش کرده‌اند که در مدل‌های درخت تصمیم، اختلافات ظریف مرتبط با تغییرات کلروفیل و آغاز کاهش سبزینه، به‌ویژه از طریق شاخص‌هایی مانند *EVI* و *GNDVI*، نقش مهمی در افزایش دقت طبقه‌بندی دارد؛ موضوعی که دقیقاً در پژوهش حاضر نیز مشاهده شد. در تاریخ ۹ شهریور، کاهش قدرت تفکیک *NDVI* و افزایش اهمیت *EVI* و شاخص‌های حساس به تغییرات کلروفیل مانند *GNDVI*، رفتار طیفی محصولات در مرحله بلوغ را به‌خوبی منعکس کرده است.

در تاریخ دوم مهر، که بخش عمده‌ای از محصولات وارد فاز پیری و کاهش رطوبت شده‌اند، نتایج با یافته‌های گائو (۱۹۹۶) و کریپین (۱۹۹۰) همخوانی دارد. این پژوهشگران نشان داده‌اند که در مراحل پایانی رشد، شاخص‌های رطوبتی مانند *NDWI* و *NMDI*، و همچنین باندهای فروسرخ کوتاه‌موج (*SWIR*)، قدرت تفکیک بیشتری نسبت به شاخص‌های مبتنی بر کلروفیل دارند. نتایج این مطالعه نیز نشان داد که *NMDI* و *SWIR1* بیشترین اهمیت را در تاریخ دوم

تغییر پارادایم، جایگزینی *SAVI* با *CI* به عنوان متغیر اصلی در تفکیک درخت، نشان می‌دهد که مدل از شاخص‌های معرف سبزینه‌گی به شاخص معرف رنگ تغییر رویکرد داده است این با وضعیت فنولوژیکی مزارع (سفیدی پنبه و زردی سایر محصولات کاملاً مطابقت دارد. استفاده از *NDVI* و *B5\_7* به مدل اجازه می‌دهد تا بین محصولات مختلف در مراحل مختلف پیری تمایز قائل شود.

این درخت تصمیم (شکل ۹) به‌منظور طبقه‌بندی چند محصول زراعی با استفاده از داده‌های سنجنش‌ازدور ماهواره لندست و بر پایه نسبت‌های بانندی و شاخص‌های طیفی طراحی شده است. در گره ریشه، شاخص *CI* به‌عنوان مؤثرترین متغیر تفکیک‌کننده به‌کار رفته که بیانگر نقش کلیدی وضعیت فیزیولوژیک و میزان سبزینه گیاه در جداسازی اولیه کلاس‌هاست. در شاخه‌ای که مقادیر *CI* پایین‌تر دارد، شاخص *NDVI* و نسبت‌های بانندی نظیر *B3\_6* قرمز به *SWIR* و *BI\_4* (آبی به قرمز) نقش تعیین‌کننده‌ای در تفکیک محصولاتی با پوشش گیاهی متوسط تا ضعیف، از جمله آفتابگردان، سویا و پنبه ایفا می‌کنند. در این بخش، پنبه عمدتاً با *NDVI* متوسط و نسبت‌های بالاتر باندهای مادون‌قرمز نزدیک به مادون‌قرمز موج کوتاه مشخص می‌شود که بازتاب‌دهنده ویژگی‌های ساختاری تاج پوشش و وضعیت رطوبتی این محصول است. در شاخه دیگر که با مقادیر بالاتر *CI* مشخص می‌شود، نسبت *(NIR/SWIR2) B5\_7* و شاخص *NDVI* به‌طور مؤثری محصولات با پوشش متراکم‌تر و رطوبت بالاتر، به‌ویژه برنج، را از سایر محصولات نظیر ذرت و پنبه تفکیک می‌کنند. به‌طور کلی، نتایج نشان می‌دهد که ترکیب شاخص‌های پوشش گیاهی با نسبت‌های بانندی حساس به رطوبت و ساختار گیاه، توان بالایی در بهبود دقت طبقه‌بندی پنبه و سایر محصولات زراعی در تصاویر لندست دارد.

نتایج به‌دست‌آمده از مدل‌سازی با الگوریتم C5 در این پژوهش نشان داد که دقت کلی طبقه‌بندی در هر سه تاریخ بررسی‌شده (۱۶ مرداد، ۹ شهریور و دوم مهر) قابل قبول بوده و الگوریتم توانسته است بخش

بالا در بلند *NIR* و بازتاب پایین در باند قرمز باعث می‌شود الگوی طیفی آن تا حد زیادی با محصولات می‌مانند آفتابگردان، ذرت و بخشی از سویا مشابه باشد. این مسئله در بسیاری از پژوهش‌ها، پیریتی و همکاران (۲۰۲۳) به‌عنوان یکی از دشوارترین بازه‌های زمانی برای تفکیک پنبه شناخته شده است، زیرا شاخص‌های کلروفیلی (*NDVI*, *EVI*, *GNDVI*) در این دوره برای اغلب محصولات الگوی مشابهی دارند. در نتیجه، دقت مدل در این تاریخ پایین‌تر از دو تاریخ دیگر است.

در ۹ شهریور، پنبه وارد مرحله گذار می‌شود؛ بخشی از تنش‌های رطوبتی و تغییرات کلروفیل آغاز شده، اما هنوز تغییرات طیفی آن به‌حدی نیست که بتواند به‌طور کامل از سایر محصولات قابل تفکیک باشد. بر اساس گزارش‌های مشابه در ادبیات پژوهشی، این مرحله معمولاً با افزایش خطاهای مدل در تفکیک پنبه همراه است، زیرا بسیاری از محصولات دیگر نیز هم‌زمان وارد فاز بلوغ می‌شوند و هم‌پوشانی طیفی در باندهای *SWIR*, *Red* و *NIR* به اوج خود می‌رسد. به همین دلیل، اگرچه دقت پنبه در این تاریخ نسبت به ۱۶ مرداد افزایش یافته است، اما همچنان با تفکیک‌پذیری کامل فاصله دارد.

اما در دوم مهر، یعنی زمانی که گیاه وارد مراحل نهایی بلوغ و باز شدن غوزه‌ها می‌شود، رفتار طیفی پنبه تفاوت دقیق‌تری با سایر محصولات پیدا می‌کند. کاهش کلروفیل، خشکی برگ‌ها، سفید شدن غوزه‌ها و افزایش روشنایی سطح مزرعه سبب می‌شود بازتاب پنبه در باندهای *SWIR* و *Red* افزایش یافته و بازتاب *NIR* کاهش یابد. این دوره زمانی از نظر طیفی، اوج جدایش‌پذیری پنبه محسوب می‌شود، زیرا بیشتر محصولات دیگر یا هنوز پوشش سبز دارند (مثل بخش‌هایی از برنج) یا الگوی زردشدگی آن‌ها در *SWIR* به شکلی متفاوت از پنبه ظاهر می‌شود (مانند آفتابگردان و سویا).

با توجه به همین ویژگی‌ها، مدل در تاریخ دوم مهر قادر بوده است حداکثر میزان تمایز را از طریق ترکیب شاخص‌های کلروفیل (*NDVI*, *EVI*, *GNDVI*) با شاخص‌های رطوبت و باندهای *SWIR* به‌دست آورد.

مهر داشته‌اند و تفکیک برنج، ذرت و سویا نسبت به دو تاریخ قبلی بهبود یافته است. این موضوع تأیید می‌کند که حساسیت بالای *SWIR* به تغییرات محتوای آب برگ، در اواخر فصل، نقش مهمی در طبقه‌بندی دقیق‌تر محصولات ایفا می‌کند؛ نتیجه‌ای که با گزارش‌های دانمز و همکاران (۲۰۲۳) در اروپا نیز همسویی کامل دارد.

همچنین دقت کلی به‌دست‌آمده در این پژوهش (در بازه ۸۲ تا ۸۴ درصد) با بسیاری از پژوهش‌های مشابه هم‌تراز است. برای نمونه، پیریتی و همکاران (۲۰۲۳) در مقایسه الگوریتم‌های مختلف یادگیری ماشین گزارش کردند که مدل‌های درخت تصمیم *C5* و *CART* معمولاً دقتی بین ۸۰ تا ۸۸ درصد در طبقه‌بندی محصولات فصل رشد ارائه می‌دهند. از سوی دیگر، مطالعات ژانگ و همکاران (۲۰۲۵) درباره نقش تاریخ تصویربرداری در بهبود طبقه‌بندی محصولات نشان داده است که بهترین نتایج معمولاً زمانی حاصل می‌شود که تاریخ تصویربرداری با مرحله رشد ویژه هر محصول همخوان باشد؛ موضوعی که در پژوهش حاضر نیز به‌خوبی مشاهده شد، به‌طوری‌که بهترین تفکیک مربوط به کلاس‌هایی بود که در تاریخ تصویربرداری در فاز رشدی متمایز قرار داشتند.

#### تعیین بهترین تاریخ شناسایی محصول پنبه:

بررسی سه تاریخ ۱۶ مرداد، ۹ شهریور و دوم مهر نشان می‌دهد که بیشترین دقت تشخیص کلاس پنبه در تاریخ دوم مهر حاصل شده است. این تاریخ در مرحله‌ای قرار دارد که فنولوژی پنبه وارد فاز نهایی فصل، شامل رسیدگی غوزه‌ها، کاهش یکنواخت کلروفیل، کاهش رطوبت برگ و نمایان شدن بخش‌های سفید غوزه‌ها می‌شود. این تغییرات سبب ایجاد امضای طیفی منحصربه‌فردی می‌شوند که در مطالعات متعدد ژانگ و همکاران (۲۰۲۵) و دانمز و همکاران (۲۰۲۳) نیز به‌عنوان بهترین فرصت زمانی برای تفکیک پنبه گزارش شده است. در جدول (۵) مقایسه ۳ مدل در طول زمان و سیر تکاملی آن ارائه شده است.

در تاریخ ۱۶ مرداد، پنبه در اوج رشد رویشی قرار دارد. غلبه پوشش سبز، مقدار بالای کلروفیل، بازتاب

مورد بررسی، بالاترین دقت تشخیص را در دوم مهر داشته باشد.

این سازگاری طیفی ناشی از وضعیت فنولوژیکی گیاه است که به طور مستقیم به افزایش دقت مدل در این تاریخ منجر شده و باعث شده پنبه در بین سه تاریخ

جدول ۵- سیر تکاملی منطق طبقه‌بندی: مقایسه ۳ مدل در طول زمان

تاریخ‌های تصویر			مشخصه مدل
دوم مهر	۹ شهریور	۱۶ مرداد	
۸۵/۱۶ درصد	۷۹/۸۸ درصد	۸۲/۵۶ درصد	دقت کلی
۹۰/۳۱ درصد	۸۶/۳۵ درصد	۸۷/۸۴ درصد	درصد پوشش (Recall) پنبه
CI	SAVI	SAVI	متغیر تفکیک کننده اصلی
NDVI, B5_7, B4_7	B3_6, B1_2, B4_7	B2_6, CI, B6_7	متغیرهای کلیدی ثانویه
۴۷	۴۸	۵۵	تعداد برگ‌ها و پیچیدگی مدل

مقایسه ماتریس‌های درهم‌ریختگی نشان می‌دهد که با حرکت به سمت تاریخ‌های منطبق‌تر با اوج رشد رویشی و تمایز فنولوژیکی محصولات، دقت کلی مدل به طور محسوسی افزایش یافته است. در هر سه تاریخ، محصولات نظیر برنج و پنبه به طور مداوم بالاترین مقادیر صحت و بازیابی را ثبت کرده‌اند که نشان‌دهنده رفتار طیفی متمایز، پوشش گیاهی متراکم‌تر و پاسخ قوی‌تر آن‌ها در شاخص‌های مبتنی بر نزدیک‌فروسرخ و رطوبت است. در مقابل، سویا در تمامی تاریخ‌ها ضعیف‌ترین عملکرد را نشان داده و کمترین میزان بازیابی را داشته است که این موضوع بیانگر هم‌پوشانی طیفی و شباهت فنولوژیکی آن با ذرت و تا حدی توتون، به‌ویژه در بازه‌های زمانی مشترک رشد، است. ذرت و آفتابگردان نیز عملکردی وابسته به زمان تصویربرداری داشته‌اند؛ به‌گونه‌ای که در تاریخ‌هایی با تفکیک فنولوژیکی بهتر، دقت این دو محصول به طور چشمگیری افزایش یافته است. علاوه بر این مدل C5 که به شکل گام به گام کدنویسی و اجراء شد بطور خودکار مهم‌ترین ویژگی‌های طیفی را در هر مرحله از رشد محصول انتخاب می‌کند. این تغییر از شاخص‌های تراکم به شاخص‌های رنگ و پیری، توانایی مدل را در درک پویایی‌های فنولوژیکی را اثبات می‌کند. به‌طور کلی، نتایج این مطالعه تأیید می‌کند که ترکیب زمان تصویربرداری مناسب با ویژگی‌های طیفی و شاخص‌های گیاهی منتخب، کلید اصلی بهبود تفکیک

تحلیل رفتار طیفی محصولات زراعی در سه تاریخ ۱۶ مرداد، ۹ شهریور و دوم مهر ۲۰۲۴ نشان می‌دهد که تغییرات فنولوژیکی، رشد و رسیدگی محصول نقش تعیین‌کننده‌ای در میزان بازتاب الکترومغناطیسی و در نتیجه در توان تفکیک مدل C5 داشته است. بر اساس الگوی بازتابی مشاهده‌شده در شکل‌های (۶)، (۷) و (۸)، تغییرات اشباع سبزینه، محتوای آب برگ و ساختار تاج پوشش مهم‌ترین عوامل مؤثر بر تغییر اهمیت باندها و شاخص‌های طیفی بوده‌اند. این یافته با گزارش‌های پژوهش‌های گذشته نیز سازگار است؛ مطالعاتی مانند علیزاده و همکاران (۲۰۱۸) و فریز و همکاران (۲۰۱۷) تأکید کرده‌اند که در طبقه‌بندی فصلی محصولات، بیشترین تمایز طیفی مربوط به تفاوت‌های محتوای کلروفیل و آب گیاه است. همچنین پیریتی و همکاران (۲۰۲۳) و فریز و همکاران (۲۰۱۷) نشان داده‌اند که عملکرد مدل‌های درخت تصمیم در طبقه‌بندی محصولات، وابستگی مستقیمی به تاریخ تصویربرداری و مرحله رشد دارد.

### نتیجه‌گیری

در یک جمع‌بندی کلی از نتایج حاصل از سه تاریخ تصویربرداری متفاوت برای تفکیک محصولات زراعی با استفاده از تصاویر ماهواره‌ای لندست و مدل درختی C5، می‌توان بیان کرد که زمان تصویربرداری نقش تعیین‌کننده‌ای در دقت و پایداری طبقه‌بندی دارد.

حاشیه مرزها نیز باشد که از محدودیت‌های این پژوهش به حساب می‌آید. با این حال، برای دستیابی به تفکیک پایدارتر محصولات دارای رفتار طیفی مشابه، به‌ویژه سویا، استفاده از داده‌های چندزمانه متراکم‌تر، شاخص‌های فنولوژیکی و یا تلفیق داده‌های سنجنده‌های با تفکیک طیفی بالاتر به‌عنوان مسیرهای پژوهشی آتی پیشنهاد می‌شود.

محصولات زراعی است و مدل C5 توانسته است به‌خوبی به تغییرات فنولوژیکی محصولات در طول فصل رشد پاسخ دهد. شایان ذکر است که برخی از خطاهای مشاهده‌شده در این پژوهش می‌تواند ناشی از عدم یکنواختی واحدهای زراعی، از جمله وجود تک‌درختان درون قطعات، کشت محصولات فرعی در میان کشت اصلی و یا اثر پیکسل‌های آمیخته در

#### منابع

1. Abdullah, A. Z., Winarno, B., and Saputro, D. R. S. 2021. The decision tree classification with C4.5 and C5.0 algorithm based on R to detect case fatality rate of dengue hemorrhagic fever in Indonesia. *Journal of Physics: Conference Series*, 1776(1), 012040.
2. Abiyat, M., Amanpour, S., Abiyat, M. and Abiyat, M. 2021. Estimation of Agricultural Cultivation Area by Landsat 8 Satellite Images (Case study: Shushtar Province). *Journal of Crop Improvement*, 24(2), 465-479.
3. Alipur, F., Aghakhani, M-H., Abbaspur-e-Fard, M-H. and Sepehr, A. 2014. Separation of area and cultivated area of agricultural products using ETM+ satellite images (Case study: Astan Quds Razavi sample farm). *Journal of Agricultural Machines*, 4(2), 244-254. (In Persian).
4. Alizadeh, P., Kamkar, B., Shataei, Sh. And Kazemi, H. 2018. Estimation of changes in land area under wheat and soybean cultivation using satellite images classification techniques in west of Golestan province. *Journal of Applied Agricultural Research*, 31(3), 41-61. (In Persian).
5. Breiman, L., Friedman, J., Olshen, R., and Stone, C. 1984. *Classification and Regression Trees*. Wadsworth.
6. Breiman, L. 2001. Random forests. *Machine Learning*, 45(1), 5-32.
7. Crippen, R.E. 1990. Calculating the vegetation index faster. *Remote Sensing of Environment*, 34(2), 71-73.
8. Dehghani, T., Ahmadpari, H. and Amini, A. 2022. 3 Assessment of land use changes using multispectral satellite images and artificial neural network. *Water and Soil Management and Modeling*, 3(2), 18-35.
9. Donmez, E., Heckeley, T., and Storm, H. 2023. Review of Satellite Remote Sensing-Based Crop Cover Classification Studies over Europe. *EarthArXiv*.
10. Friesz, A. M., Wylie, B. K., and Howard, D. M. 2017. Temporal expansion of annual crop classification layers for the CONUS using the C5 decision tree classifier. *Remote Sensing Letters*, 8(5), 389-398.
11. GAO, B.C. 1996. NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment J*, 58(3), 257-266.
12. Guo, Z., Shi, Y., Huang, F., Fan, X., and Huang, J. 2021. Landslide susceptibility zonation method based on C5.0 decision tree and K-means cluster algorithms to improve the efficiency of risk management. *Geoscience Frontiers*, 12(6), 101249.
13. Han, J., Kamber, M., and Pei, J. 2012. *Data Mining: Concepts and Techniques*.
14. Huete, A.R. 1988. A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment J*, 25(2), 295-309.
15. Huete, A.D. 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment J*, 83(3), 195-213.
16. Jensen, J. R. 2016. *Introductory Digital Image Processing. A Remote Sensing Perspective*. Jordan, C.F. 1969. Derivation of leaf area index from quality of light on the forest floor. *Ecology*, 50, 663-666.

17. Kamkar, B., Alizadeh Dehkordi, P., Aalae Bazkiaee, P. and Abdi, O. 2021. Determination of the compliance of soybean lands with land suitability maps (Case Study: Golestan Province). *Agricultural Engineering (Scientific Journal of Agriculture)*, 44 (1), 124-137.
18. Momtazi Burojeni, M. and Sarmadian, F. 2022. Spatial prediction of soil classes using C5.0 boosted decision tree model in part of Abyek lands. *Journal of range and watershed management*, 75(4), 553-572.
19. Preethi, S., Ramalingam, K., Pazhanivelan, S., Muthumanickam, D., Kaliaperumal, R., and Moorthi, N. R. 2023. Comparing the Effectiveness of Different Machine Learning Algorithms for Crop Cover Classification Using Sentinel 2. *International Journal of Environment and Climate Change*, 13(10), 571–582.
20. Quinlan, J. R. 1993. *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers Inc.
21. Quinlan, J. R. 1996. Improved use of continuous attributes in C5.0. *Journal of Artificial Intelligence Research*.
22. Rondeaux, G., Steven, M. and Baret, F. 1996. Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment J*, 55, 95–107.
23. Rouse, J.W., Hass, R.H., Schell, J.A. and Deering, D.W. 1974. Monitoring Vegetation Systems in the Great Plains with ERTS, *Third ERTS symposium, NASA SP-351(I)*, 309-317.
24. Shannon, C. E. 1948. A mathematical theory of communication. *Bell System Technical Journal*
25. Smith, A. M., Bourgeois, G., Teillet, P. M., Freemantle, J. and Nadeau, C. 2008. A comparison of NDVI and NDVI2 for estimating LAI using CHRIS imagery: a case study in wheat. *Canadian Journal of Remote Sensing*, 34(6), 539-548.
26. Song, C., Woodcock, C. E., Seto, K. C., Lenney, M. P. and Macomber, S. A. 2001. Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? *Remote Sensing of Environment*, 75(2), 230–244.
27. Steven, M.D. 1998. The Sensitivity of the OSAVI Vegetation Index to Observational Parameters. *Remote Sensing of Environment J*, 63, 49–60.
28. Wang, L. and Qu, J.J. 2007. NMDI: A normalized multi-band drought index for monitoring soil and vegetation moisture with satellite remote sensing. *Geophysical Research Letters*, 34(20).
29. Wang, T. 2024. Improved random forest classification model combined with C5.0 algorithm for vegetation feature analysis in non-agricultural environments. *Scientific Reports*.
30. Zhang, C., Kerner, H., Wang, S., Hao, P., Li, Z., Hunt, K. A., Abernethy, J., Zhao, H., Gao, F., Di, L., Guo, C., Liu, Z., Yang, Z., Mueller, R., Boryan, C., Chen, Q., Beeson, P. C., Zhang, H. K., and Shen, Y. 2025. Remote sensing for crop mapping: A perspective on current and future crop-specific land cover data products. *Remote Sensing of Environment*, 330, 114995.
31. Zhao, J., Zhang, D., Wang, D., and Huang, W. 2011. Integrating Landsat TM Imagery and See5 Decision-Tree Software for Identifying Croplands: A Case Study in Shunyi District, Beijing. In Z. Gong et al. (Eds.), *Wireless, Information, and Systems Modeling*, 251–258.