



علوم محیطی

علوم محیطی سال هشتم، شماره سوم، بهار ۱۳۹۰
ENVIRONMENTAL SCIENCES Vol.8, No.3, Spring 2011

۱-۲۰

توسعه مدل مفهومی مبتنی بر الگوریتم Decision Tree و داده کاوی به منظور آشکارسازی تغییرات پوشش اراضی با استفاده از تصاویر سنجنده‌ی TM و داده‌های کمکی

«مطالعه موردی: بخش مرکزی شهرستان بویراحمده»

حمیدرضا جعفری^۱، محمد حمزه^۲، حسین نصیری^۳، یوسف رفیعی^{۳*}

۱- دانشیار گروه برنامه‌ریزی و مدیریت محیط زیست، دانشگاه تهران

۲- دانشجوی کارشناسی ارشد سنجش از دور و سیستم اطلاعات جغرافیایی، دانشگاه تهران

۳- دانشجوی کارشناسی ارشد برنامه‌ریزی و مدیریت محیط زیست، دانشگاه تهران

تاریخ پذیرش: ۸۹/۱۱/۱۸

تاریخ دریافت: ۸۹/۷/۴

Developing Decision Tree and Data Mining Based Conceptual Model for Detecting Land Cover Changes Using TM Images and Ancillary Data Study Area: Central Section of Boyerahmad County

Hamid Reza Jafari,¹ Mohammad Hamzeh,² Hossein Nasiri² and Yusef Rafii^{3*}

1. Assoc. Professor, Department of Environmental Planning and Management, Faculty of Environment, University of Tehran, Tehran-Iran

2. MSc. Student, Department of Cartography (RS & GIS), University of Tehran

3. MSc. Student, Department of Environmental Planning and Management, University of Tehran

Abstract

Rapid urban growth and industrialization have caused many environmental problems in a number of cities around the world. Knowledge about land cover/land use changes in the long term is very important for urban managers and policy makers in order to evaluate and predict the resulting problems. Remote sensing is an effective tool for monitoring these changes in urban areas and its periphery. Over recent decades, Yasouj City has developed and affected its surrounding environment due to urban growth and immigration. The objective of this research is to develop a Decision Tree and data mining based conceptual model for land cover change detection using a Landsat Thematic Mapper (TM) and ancillary data in the central Section of Boyerahmad County from 1990 to 2009. Based on findings of the study, the overall six-class classification accuracies for 1990 and 2009 were, respectively, 93.16% and 93.37%. The overall accuracy of land cover change maps, generated from post-classification change detection methods and evaluated using two approaches, ranged from 85.6% to 86.98%. The maps also showed that between 1990 and 2009 the urban area increased by approximately 19.28% while agriculture and forest decreased by 31.76% and 7.32% respectively.

Keywords: Change detection, Conceptual model, Data mining, Decision Tree, TM, Boyerahmad County.

چکیده

رشد سریع شهرها و روندهای صنعتی شدن، مشکلات زیست‌محیطی زیادی را در بسیاری از شهرهای جهان ایجاد کرده است. آگاهی از روندهای تغییرات پوشش و کاربری اراضی طی دوره‌های زمانی طولانی برای برنامه‌ریزان و مدیران به منظور ارزیابی و پیش‌بینی مشکلات ناشی از این تغییرات بسیار بااهمیت است. سنجش از دور ابزار موثری برای پایش تغییرات پوشش اراضی در مناطق شهری و حوالی آن می‌باشد. شهر یاسوج در چند دهه اخیر بواسطه رشد و بدنبال مهاجرت افراد، گسترش زیادی پیدا نموده و اثرات فراوانی بر محیط‌زیست پیرامون خود بر جای گذاشته است. هدف از این پژوهش، توسعه یک مدل مفهومی مبتنی بر الگوریتم Decision Tree و داده کاوی برای آشکارسازی تغییرات پوشش اراضی با استفاده از تصاویر سنجنده TM در محدوده بخش مرکزی شهرستان بویراحمده برای سال‌های ۱۹۹۰ تا ۲۰۰۹ می‌باشد. براساس یافته‌های این تحقیق، صحت کلی طبقه‌بندی برای سال‌های ۱۹۹۰ و ۲۰۰۹ به ترتیب ۹۳/۱۶ و ۹۳/۳۷ درصد می‌باشد. صحت کلی نقشه تغییرات پوشش اراضی نیز به روش مقایسه پس از طبقه‌بندی با استفاده از دو تکنیک محاسبه گردید که نتایج ۸۵/۶ تا ۸۶/۹۸ درصد حاصل شد. هم‌چنین نقشه تغییرات نشان داد که از سال ۱۹۹۰ تا ۲۰۰۹ مساحت منطقه مسکونی حدود ۱۹/۲۸ کیلومتر مربع افزایش یافته و اراضی کشاورزی، جنگلی به ترتیب کاهشی ۳۱/۷۶ و ۷/۳۲ درصدی داشته‌اند.

کلید واژه‌ها: آشکارسازی تغییرات، مدل مفهومی، داده کاوی، Decision Tree، TM، بویراحمده.

* Corresponding author. E-mail Address: yusefrafii@ut.ac.ir

مقدمه

امروزه رشد بی‌رویه جمعیت، صنعتی شدن شهرها، اعمال سیاست‌های اقتصادی ویژه از سوی دولت‌ها و مهاجرت روستائیان به شهرها سبب گسترش سریع مناطق شهری شده است. رشد منطقه شهری بویژه حرکت مناطق مسکونی و تجاری به سمت مناطق روستایی و کشاورزی در حاشیه مناطق شهری به عنوان نشانه‌ای از حیات اقتصادی ناحیه مد نظر قرار می‌گیرد؛ اما این رشد به طور فزاینده‌ای اثرات سوء زیست‌محیطی شامل کاهش کیفیت آب و هوا، از بین رفتن اراضی جنگلی، تبدیل زمین‌های کشاورزی به مناطق مسکونی یا اراضی بایر و در نتیجه کاهش تولید محصولات کشاورزی، فرسایش خاک، افزایش سیلاب، افزایش خطر انقراض حیات وحش و تأثیرات اجتماعی-اقتصادی اعم از چندپارگی اجتماعی و هزینه‌های زیرساختی را ایجاد می‌نماید. تغییرات اراضی در حاشیه شهرها به گسترش سریع حومه‌های کم جمعیت به نواحی روستایی و کشاورزی مربوط می‌گردد. این تغییرات و تأثیرات ناشی از آن باید توسط مدیران و سیاست‌گذاران ناحیه‌ای و محلی مورد توجه قرار گیرد تا تصمیم‌گیری‌های صحیح و آگاهانه‌ای اتخاذ گردد که جنبه‌های مثبت و تأثیرات منفی توسعه و رشد شهر را به منظور حفظ منابع محیطی و افزایش رفاه و آسایش اجتماعی-اقتصادی به‌طور موثری متعادل کنند. لزوم حفظ محیط‌زیست و بهره‌برداری بهینه از منابع طبیعی، از جمله ضرورت‌های توسعه پایدار محسوب می‌شود.

یاسوج از جمله شهرهایی است که در چند دهه اخیر بواسطه رشد و مهاجرت افراد به سوی مناطق شهری، گسترش بسیاری یافته و اثرات زیست‌محیطی فراوانی بر محیط پیرامون خود بر جای گذاشته است. در راستای شناخت و برنامه‌ریزی برای این امر، پژوهش حاضر به توسعه یک مدل مفهومی و استخراج موثر و دقیق اطلاعات بر مبنای این مدل جهت آشکارسازی تغییرات انواع پوشش اراضی می‌پردازد تا بدین طریق کمکی به

برنامه‌ریزان جهت ارزیابی و پیش‌بینی مشکلات ناشی از این تغییرات باشد.

با توجه به نیاز مداوم تصمیم‌گیرندگان و سیاست‌گذاران به اطلاعات مکانی به روز در مورد الگوها و روندهای تغییرات پوشش و کاربری اراضی، تحقیقاتی در مورد پتانسیل داده‌های ماهواره‌ای برای آشکارسازی یا پایش تغییرات کاربری و پوشش اراضی صورت پذیرفته است که در ذیل به پاره‌ای از آن‌ها اشاره می‌شود:

Alberti و همکاران (2004)، Goetz و همکاران (2004) و Yang (2002) نشان دادند که سنجش از دور ماهواره‌ای برای تولید اطلاعات مکانی دقیق و به هنگام در مورد تغییرات پوشش و کاربری اراضی شهری پتانسیل کافی را دارد. اگرچه تغییرات کاربری و پوشش اراضی بوسیله برخی روش‌های سنتی نیز می‌تواند تعیین گردد ولی سنجش از دور ماهواره‌ای اطلاعات مفیدی را در مورد توزیع مکانی کاربری و پوشش اراضی و تغییرات آن‌ها همراه با مزایایی مثل صرفه‌جویی در هزینه و زمان مهیا می‌کند.

Kemper و همکاران (2001) به بررسی تغییرات کاربری اراضی شهر در یک دوره ۵۰ ساله در استانبول پرداختند. در این مطالعه شش کلاس کاربری اراضی انتخاب و برای ۴ دوره آنالیز شدند که برای این منظور از تصاویر IRS، IKONOS و عکس‌های هوایی و روش تفسیر چشمی استفاده گردید و تغییرات موجود در بین این کاربری‌ها در این ۵ دهه مورد بررسی و تخمین قرار گرفت. Kaya و همکاران (2006) به پایش رشد شهر در قسمتهای اروپایی استانبول پرداختند. هدف این تحقیق کمی‌سازی رشد شهر در قسمت‌های اروپایی استانبول در دو دهه اخیر بود. شش نوع پوشش اراضی با استفاده از تصاویر ماهواره لندست برای سال‌های ۱۹۸۷، ۱۹۹۲، ۱۹۹۷ و ۲۰۰۱ در این پژوهش مشخص شد و از تغییرات پوشش اراضی در طی این دوره‌ها برای تعیین سرعت تغییرات استفاده گردید.

- توسعه یک مدل مفهومی مبتنی بر الگوریتم Decision Tree و داده کاوی جهت بالا بردن دقت طبقه‌بندی تصاویر ماهواره‌ای،
- ارزیابی صحت طبقه‌بندی‌ها تصاویر چندزمانه سنجنده TM، برآورد صحت آشکارسازی تغییرات و تحلیل تغییرات پوشش اراضی.

مواد و روش‌ها

منطقه مورد مطالعه

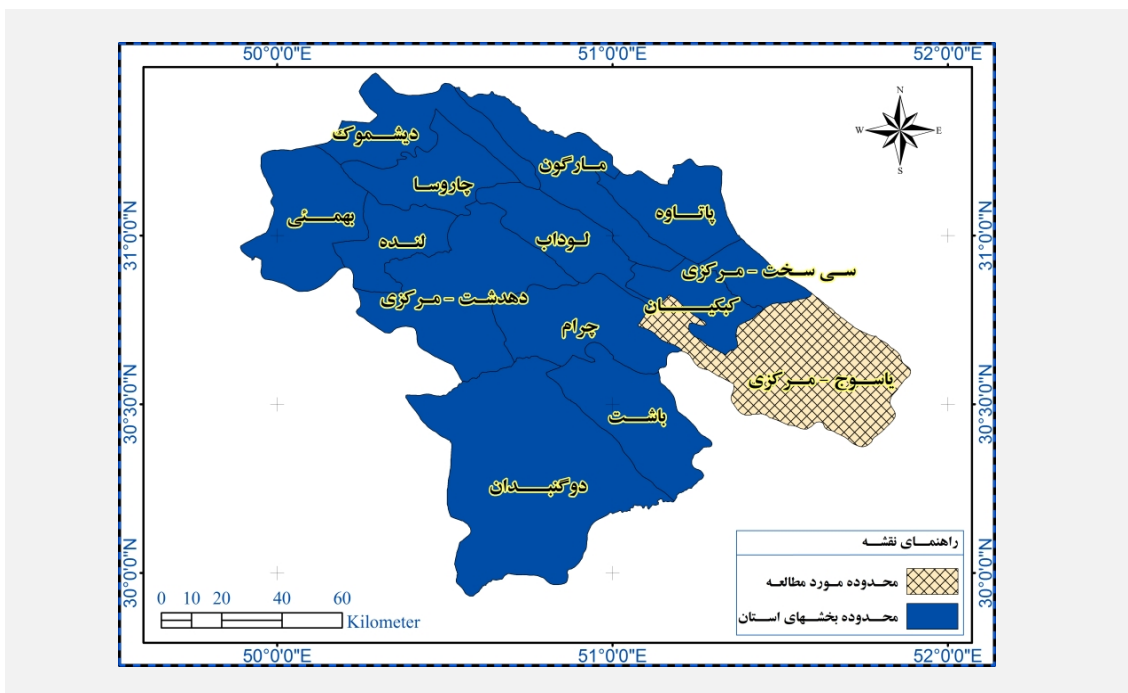
شهرستان بویراحمد یکی از شهرستان‌های استان کهگیلویه و بویراحمد است، که در شرق این استان واقع شده است. مرکز این شهرستان، شهر یاسوج است. این شهرستان از شمال به استان اصفهان، از شرق و جنوب به استان فارس، از غرب به شهرستان کهگیلویه و بویراحمد (به مرکزیت یاسوج) واقع در استان کهگیلویه و بویراحمد می‌باشد که در طول ۵۱ درجه و ۳۶ دقیقه و عرض ۳۰ درجه و ۴۰ دقیقه قرار گرفته و ارتفاع متوسط آن از سطح آب‌های آزاد حدود ۱۸۵۰ متر است (شکل ۱).

Alavi Panah (2005) به تشخیص تغییرات و بیابان‌زایی با استفاده از تصاویر سنجنده TM، ETM+ و MSS از طریق تفریق تصاویر و مقایسه پس از طبقه‌بندی پرداخته است. نتایج حاصله نشانگر توانایی داده‌های ذکر شده برای تهیه نقشه‌های تغییرات است.

Feyzi zade و Haj mirrahimi (2007) از تصاویر سنجنده TM ماهواره لندست و HDR ماهواره اسپات استفاده نموده و تغییرات فضای سبز شهر تبریز را با استفاده از روش طبقه‌بندی شیء گرا مورد مطالعه قرار داده‌اند. بر اساس این تحقیق مشخص شد که بیش از ۴۶ درصد فضای سبز شهر تبریز در یک دوره زمانی ۱۶ ساله تخریب شده است.

Seto (2002) با استفاده از تصاویر سنجنده TM نقشه کاربری‌های اراضی دلتای رودخانه بی آرل واقع در قسمت جنوبی کشور چین را استخراج نمود و با استفاده از همین تصاویر چند زمانه تغییرات کاربری اراضی را در این منطقه مشخص نمود.

حال با توجه به مطالب بیان شده، اهداف پژوهش حاضر جهت آشکارسازی تغییرات پوشش اراضی بخش مرکزی شهرستان بویراحمد عبارتند از:



شکل ۱ - موقعیت جغرافیایی ناحیه مورد مطالعه

داده‌های مورد استفاده

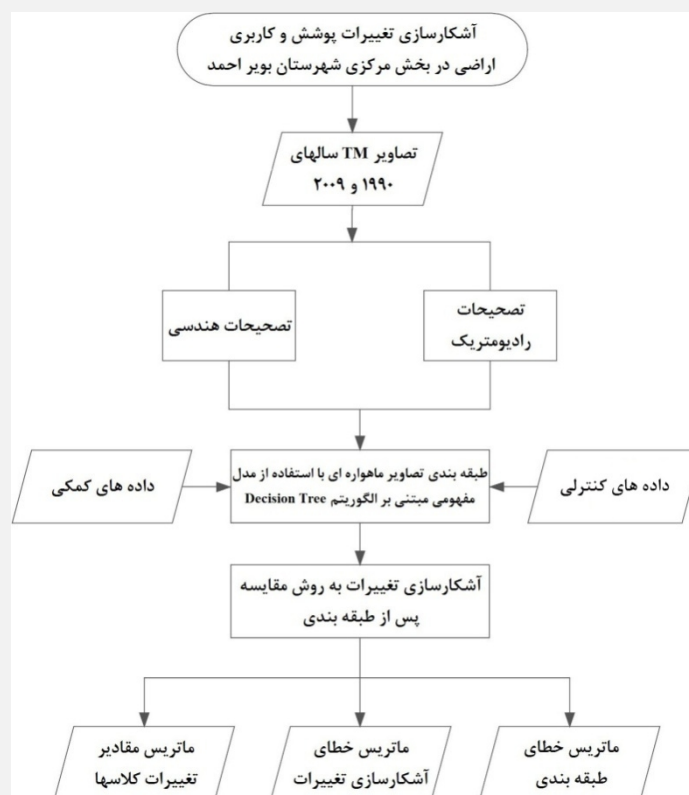
در مجموع دو سری داده در این مطالعه مورد استفاده قرار گرفته است. اول دو تصویر سنجنده TM ماهواره لندست است که به ترتیب در تاریخ‌های ۱۹۹۰/۰۶/۰۶ و ۲۰۰۹/۰۵/۲۵ اخذ شده و دوم داده‌های کمکی یعنی مدل رقومی ارتفاع مستخرج از نقشه توپوگرافی ۱:۲۵۰۰۰ سازمان نقشه‌برداری است (جدول ۱).

روش‌شناسی

برای آشکارسازی تغییرات پوشش اراضی عملیات پردازش بر روی تصاویر ماهواره‌ای در سه مرحله پیش پردازش، پردازش و پس پردازش انجام شد. بطور کلی فرایند انجام این تحقیق مبتنی بر مراحل اصلی زیر است که بصورت شماتیک در شکل ۲ نشان داده شده است که در ادامه به تشریح آن‌ها پرداخته خواهد شد.

جدول ۱- مشخصات داده‌های مورد استفاده

نوع	تاریخ	شماره	شماره	شماره	قدرت تفکیک
TM	۱۹۹۰/۰۶/۰۶	۱ تا ۷	۱۶۳	۳۹	۸ بیت ۳۰ متر
TM	۲۰۰۹/۰۵/۲۵	۱ تا ۷	۱۶۳	۳۹	۸ بیت ۳۰ متر
DEM	-	-	-	-	۱۰ متر



شکل ۲- فرایند آشکارسازی تغییرات پوشش و کاربری اراضی

پیش پردازش

در مرحله پیش‌پردازش، یک سری عملیات روی داده‌های خام (قبل از هر گونه پردازش بصری یا رقومی)، در ارتباط با خطاهای رادیومتری و هندسی انجام می‌گیرد. از این رو با توجه به هدف مطالعه که آشکارسازی تغییرات با استفاده از این دو تصویر می‌باشد، در ابتدا اقدام به اعمال تصحیحات مورد نیاز در مرحله‌ی پیش‌پردازش شد. در این مرحله تصحیحاتی که بر روی تصاویر ماهواره‌ای صورت گرفت عبارتند از:

- تصحیحات هندسی (Geometric correction)،
 - تصحیحات رادیومتریک (Radiometric correction).
- در ابتدای امر از سری تصحیحات هندسی عملیات زمین مرجع کردن بر روی تصاویر اعمال گردید. برای انجام این کار ۲۱ نقطه کنترل زمینی با پراکنش مناسب با استفاده از سیستم موقعیت‌یاب جهانی از سطح منطقه جمع‌آوری شد تا اینکه عملیات زمین مرجع کردن با دقت بهتری صورت گیرد. برای تبدیل مختصات تصویر چند جمله‌ای درجه دوم به کار گرفته شد. ابتدا تصویر سال ۲۰۰۹ با استفاده از این نقاط کنترل زمینی، زمین مرجع گردید، سپس به منظور تصحیح هندسی تصویر دوم (سال ۱۹۹۰) از روش ثبت تصویر به تصویر با ۳۳ نقطه کنترلی استفاده گردید. در این روش، تصویر سال ۲۰۰۹ که در مرحله قبل تصحیح شده بود به عنوان مرجع در نظر گرفته شده و مختصات تصویر سال ۱۹۹۰ با استفاده از نقاط کنترلی متناظر انتخاب شده بر روی هر دو تصویر بر سیستم مختصات تصویر اول منطبق گردید. برای تعیین ارزش رقومی پیکسل‌های تصویر تصحیح شده در هر دو مرحله از روش نمونه‌گیری مجدد نزدیک‌ترین همسایه استفاده شد و تصاویر با RMSE معادل ۰/۳۵ پیکسل، زمین مرجع گردیدند. در گام بعد اثرات اتمسفری و اثرات توپوگرافی و تابع توزیع انعکاس

دوجهته^۱ بر بازتاب طیفی نیز با استفاده از مدل ATCOR3 اصلاح گردید. الگوریتم ATCOR3 تا حد زیادی اثرات جوی، اثرات سایه و روشنایی^۲ را در تصویربرداری طیفی با دخالت دادن مدل رقومی زمین و شیب کاهش می‌دهد. در مرحله‌ی بعد جهت تعیین کلاس‌ها از روش طبقه‌بندی Decision Tree و تکنیک‌های داده‌کاوی استفاده شد.

پردازش

طبقه‌بندی تصویر به روش (DT) Decision Tree

نظر به اینکه در این تحقیق برای طبقه‌بندی تصاویر یک مدل مفهومی به کار گرفته شده که این مدل بر مبنای روش DT است از این رو به تشریح آن می‌پردازیم.

DT برای کاربردهای سنجش از دوری یوسيله متخصصین این رشته در دهه ۱۹۷۰ ارزیابی شده است (Swain and Hauska, 1977). در سال‌های اخیر این روش به صورت تدریجی از کاربردهای تجاری به سمت علوم طبیعی وارد شده و در طبقه‌بندی پوشش اراضی، عملکرد موقیت‌آمیزی داشته است (Hansen et al., 1996 ; Brodley et al., 1999 ; Vogelmann et al., 2001 ; Lawrence, Wrigh, 2001).

در اغلب روش‌های طبقه‌بندی تک‌مرحله‌ای تصاویر ماهواره‌ای، مانند روش حداکثر شباهت^۳ یا حداقل فاصله^۴ در مورد هر پیکسل فقط یک تصمیم اتخاذ می‌گردد، در نتیجه‌ی این تصمیم پیکسل به یکی از کلاس‌های موجود تعلق می‌گیرد. در کنار این روش‌ها، تکنیک‌های طبقه‌بندی چندمرحله‌ای هم امکان‌پذیر هستند؛ بدین ترتیب که مجموعه‌ای از تصمیمات اتخاذ می‌شوند تا پیکسل به درستی طبقه‌بندی شود. معمول‌ترین روش‌های طبقه‌بندی چندمرحله‌ای، Decision Trees نامیده می‌شوند. این دسته، از طبقه‌بندی کننده‌های^۵ مرتبط با هم تشکیل می‌شوند که هیچکدام به تنهایی نمی‌توانند امر

طبقه‌بندی تصویر را بطور کامل انجام دهند؛ بلکه هر مؤلفه طبقه‌بندی کننده، تنها قسمتی از کار را انجام می‌دهد. در واقع می‌توان گفت که یک DT، نمایشی از شاخه‌ها و گره‌هاست. هر گره به مجموعه‌ای از پاسخ‌های ممکن منتهی می‌گردد که مورد یا داده وارد شده به گره را به زیرمجموعه‌های متناظر با نتایج متفاوت تقسیم می‌کند. ساده‌ترین نوع از DT ها نوع شاخه دودویی است که در هر گره یا مؤلفه طبقه‌بندی کننده آن، داده مربوط به تصویر به یکی از دو کلاس ممکن یا یکی از دو گروه کلاس موجود تعلق می‌یابد. خیلی اوقات تکنیک‌های DT می‌توانند به صورت غیر خود کار طراحی شوند. بخصوص هنگامی که برای انجام طبقه‌بندی خیلی خاصی به کار گرفته می‌شوند (Swain and Hauska, 1977)، ولی این روشها می‌توانند بصورت خود کار نیز طراحی شوند. سه امر مهم در طراحی روش DT وجود دارد:

اول یافتن ساختار بهینه برای شاخه، دوم انتخاب زیرمجموعه مطلوب خصوصیات در هر گره و مورد آخر انتخاب قاعده تصمیم برای استفاده در هر گره (Richards and Jia, 2006). یک ساختار شاخه‌ای بهینه می‌تواند با کم‌ترین نرخ خطا یا حداقل تعداد گره‌ها فرض شود. هم چنین باید به اشتراک کلاس‌ها و اینکه چه تعداد شاخه و لایه مورد استفاده قرار می‌گیرند توجه کافی مبذول داشت. از آنجایی که تعداد ساختارهای شاخه‌ای ممکن، حتی برای تعداد نسبتاً کمی از کلاس‌ها خیلی زیاد است، طراحی یک طبقه‌بندی کننده مطلوب بسیار دشوار خواهد بود (Mui and Fu, 1980).

کارایی و صحت طبقه‌بندی شدیداً به انتخاب شاخه‌ها بستگی دارد. از این رو روش‌های ابتکاری متعددی برای طراحی DT بکار گرفته شده است که هر کدام از آن‌ها عیوب و مزایایی دارند. برای آسان‌تر نمودن کار طراحی شاخه‌های تصمیم‌گیری، غالباً روش دودویی استفاده

می‌شود. توانایی جداسازی کلاس‌ها ضرورتاً با به کارگیری روش دودویی تضعیف نمی‌شود، چون یک DT کلی می‌تواند به شکل مناسبی به شاخه‌های دودویی معادل با آن تبدیل شود، بدون اینکه کاهش دقتی صورت پذیرد (Rounds, 1980). بصورت کلی در روش DT، دانش تحلیل‌گر در ساختار بندی شاخه‌ها نقش موثری دارد. کاربر می‌تواند در روش DT از الگوریتم‌های دیگری همچون حداکثر شباهت، حداقل فاصله و موارد مشابه هم استفاده کند، ولی بهتر است این نوع الگوریتم‌ها زمانی بکار گرفته شوند که نیاز به جداسازی کلاس‌های کوچک وجود دارد. علاوه بر این، داده‌های مرتبط با توپوگرافی مانند ارتفاع، مقدار شیب، جهت شیب و حتی داده‌های اقلیمی مانند دما و ... می‌توانند در روش DT دخالت داده شوند و طبقه‌بندی را بهبود بخشند.

مدل مفهومی مبتنی بر الگوریتم Decision Tree

از آنجایی که در این پژوهش هدف شناسایی انواع پوشش اراضی است، یک مدل مفهومی برای این منظور توسعه داده شده و استخراج موثر و دقیق اطلاعات بر مبنای این مدل صورت گرفت. بر اساس باز دیدهای میدانی صورت گرفته پوشش کلی اراضی در ناحیه مورد مطالعه شامل اراضی صخره‌ای، بایر، آب، پوشش گیاهی و اراضی مسکون شهری می‌باشد. مدل مفهومی ارائه شده در این تحقیق به این صورت عمل می‌کند که به صورت شاخه‌ای در هر قسمت یک پوشش خاص از اطلاعات موجود در تصویر را استخراج و جدا می‌کند. در هر نقطه انشعابی یا گره مربوط به DT با استفاده از قواعد یا شاخص‌های تعریف شده، پوشش خاصی از اطلاعات تصویر شناسایی و مجزا می‌گردد. در این مدل از شاخص‌های آب، ساخت و ساز انسانی، شاخص‌های گیاهی، دمای سطحی زمین، شیب و مقادیر انعکاس

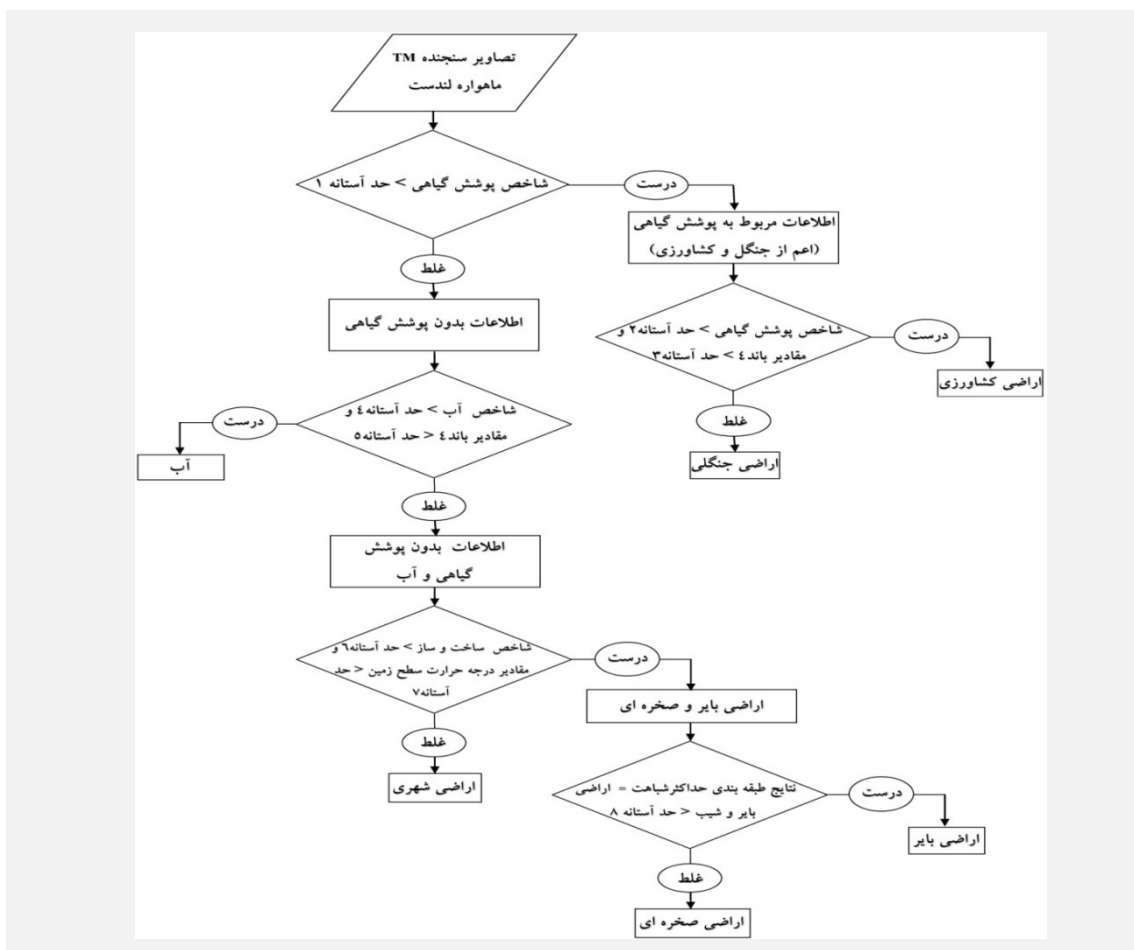
ارزیابی و بررسی پوشش‌های گیاهی با استفاده از داده‌های چندطیفی ماهواره‌ای به کار گرفته می‌شوند (Fatemi and rezaii, 2005). بیشترین باندهایی که در محاسبه شاخص‌های گیاهی مورد استفاده قرار می‌گیرند در محدوده باندهای قرمز و مادون قرمز نزدیک هستند. دلیل این امر خاصیت جذبی رنگدانه‌های گیاهی از جمله کلروفیل است که باعث انعکاس کم‌تر گیاهان در باند قرمز و انعکاس شدید آن‌ها در بخش مادون قرمز طیف الکترومغناطیس است. در این پژوهش برای تفکیک پوشش گیاهی از سایر اطلاعات و کلاس اراضی کشاورزی از جنگلی به ترتیب از دو شاخص $^{\vee}$ NDVI و $^{\wedge}$ SAVI استفاده گردید که به توضیح آن‌ها پرداخته می‌شود.

سطحی مربوط به باندها استفاده شده است. مدل مفهومی طرح شده بر روی داده‌های دو زمانه ماهواره لندست اعمال گردید. شکل ۳ چارچوب کلی مدل را به صورت یک روندنما نشان می‌دهد.

حال ابتدا به توضیح شاخص‌ها یا قواعد استفاده شده در مدل پرداخته می‌شود و سپس در ادامه در مورد نحوه کارکرد مدل مفهومی طراحی شده توضیحاتی ارائه می‌گردد.

شاخص‌های پوشش گیاهی

شاخص‌های پوشش گیاهی تبدیلات ریاضی هستند که براساس باندهای مختلف سنجنده‌ها تعریف شده و برای



شکل ۳- مدل مفهومی مبتنی بر الگوریتم Decision Tree

شاخص تفاضلی نرمال شده پوشش گیاهی (NDVI)

این شاخص یکی از پرکاربردترین و رایج‌ترین شاخص‌های مربوط به پوشش گیاهی می‌باشد که برحسب دو باند قرمز و مادون قرمز نزدیک به صورت زیر تعریف می‌شود.

$$NDVI = (NIR - RED) / (NIR + RED)$$

در این شاخص فرض بر این است که ناحیه تحت بررسی دارای نوع خاک یکسانی می‌باشد، بنابراین اگر این فرض صادق نباشد ناهمسانی انواع خاک بر روی نتایج حاصل از NDVI تاثیر خواهد گذاشت و نتایج نادرستی حاصل خواهد شد. این شاخص دارای مقادیر نرمال در دامنه بین -1 و +1 است که بررسی و نمایش مقادیر را براحتی امکان‌پذیر می‌سازد. مقادیر مختلف NDVI می‌تواند نمایانگر پوشش‌های متفاوت باشد.

شاخص پوشش گیاهی تعدیل کننده اثر خاک (SAVI)

این شاخص گیاهی با وارد کردن یک پارامتر به فرمول NDVI سعی در کاهش اثر پس زمینه‌های حاکی در مناطق داری پوشش گیاهی دارد. Huete (1988)، SAVI را برای استخراج اطلاعات پوشش گیاهی پیشنهاد کرد که به صورت زیر فرمول‌بندی شده است.

$$SAVI = [(NIR - RED) (1 + L)] / (NIR + RED + L)$$

که L فاکتور تعدیل کننده اثر خاک است که معمولاً با سعی و خطا محاسبه می‌شود. مقدار L برای مناطق با پوشش گیاهی پایین 1 و برای مناطق با پوشش گیاهی متوسط 0.5 در نظر گرفته می‌شود، که این مقدار با توجه به عدم اطلاعات کافی از مناطق مورد مطالعه معمولاً به عنوان پیش فرض استفاده می‌شود. دامنه مقادیر SAVI بین -1 و +1 محدود می‌شود. نقطه قوت این شاخص

گیاهی نسبت به NDVI، حذف اثر پس زمینه توسط عامل L است. در این تحقیق با توجه به درصد پوشش گیاهی محاسبه شده از شاخص NDVI در منطقه مورد مطالعه مقدار 0.5 برای L به منظور کاهش اثر پس-زمینه‌های خاک استفاده شد.

شاخص آب

در این پژوهش برای جداسازی کلاس آب از شاخص آب استفاده گردید. آب در محدوده سبز طیف الکترومغناطیس انعکاس بیشتری نسبت به محدوده مادون قرمز نزدیک دارد. بر همین اساس Mcfeeters (1996) شاخص تفاضلی نرمال شده آب (NDWI) را به صورت فرمول زیر ارائه کرده است.

$$NDWI = (GREEN - NIR) / (GREEN + NIR)$$

که Green مقادیر باند سبز می باشد.

شاخص ساخت و ساز انسانی

برای جداسازی خاک از عوارض ساخته دست بشر مانند مناطق مسکونی، زمانی که سایر عوارض مشکل‌زا (مانند عوارض مربوط به کشاورزی) از تصویر حذف شده‌اند، می‌توان از بافت⁹ تصویر استفاده کرد. از ویژگی‌های بافت می‌توان به کنتراست، آنتروپی¹⁰ و همبستگی و غیره اشاره کرد (Jiang et al., 2004). اما این روش فقط برای تصاویری با قدرت تفکیک مکانی بالا مانند SPOT می‌تواند مفید باشد. نتایج کار سایر محققین حاکی از دقت کم روش استفاده از بافت برای جداسازی خاک از سایر عوارض با استفاده از تصاویر سنجنده TM می‌باشد. تصاویری با قدرت تفکیک مکانی پایین معمولاً اطلاعات طیفی بیشتری دارند. به همین دلیل ما سعی کردیم برای حل این مشکل راه‌حلی از جنبه طیفی بیابیم. Yong (2003) NDBI¹¹ را بر اساس این اصل که مناطق دارای ساخت و ساز انسانی دارای انعکاس بالاتری در

محدوده مادون قرمز میانی نسبت به مادون قرمز نزدیک هستند ارائه کرد که فرمول آن بصورت زیر است.

$$NDBI = (MIR - NIR) / (MIR + NIR)$$

که MIR مقادیر باند مادون قرمز میانی است. این شاخص می تواند بخشی از اطلاعات مربوط به مناطقی با ساخت و ساز انسانی را استخراج نماید ولی باید در نظر داشت که این شاخص به صورت نسبی دقت کمتری نسبت به شاخص های پوشش گیاهی (مانند NDVI) دارد. به سخن دیگر بعد از اعمال این شاخص، نتایج استخراج شده ممکن است با سایر کلاس ها اعم از پوشش گیاهی یا آب مخلوط باشد. بنابراین برای حل این مشکل ما پس از جداسازی پوشش گیاهی و آب شاخص را بر روی اطلاعات فاقد پوشش گیاهی و آب اعمال کردیم.

عامل دمای سطح زمین^{۱۲}

با توجه به درجه حرارت بالای مناطق مسکونی نسبت به پوشش های دیگر و دقت نسبتاً پایین NDBI برای جداسازی مناطق مسکونی از اطلاعات فاقد پوشش گیاهی و آب، علاوه بر NDBI به صورت اشتراکی از درجه حرارت سطح زمین هم برای جداسازی مناطق مسکونی از اطلاعات فاقد پوشش گیاهی و آب، استفاده شد. نقشه حرارت سطح زمین با استفاده از باند ۶ سنجنده TM ماهواره لندست و روش کانال منفرد-Jime'nez و Mun'oz و Sobrino (2003) با برنامه نویسی در محیط IDL تهیه گردید. ضمناً برای محاسبه توان تشعشع سطح زمین^{۱۳} هم از مقادیر NDVI استفاده شد (Alavi panah *et al.*, 2006). علاوه بر شاخص های فوق همان گونه که در روندنما مشاهده می شود برای جداسازی برخی پوشش های سطح زمین از شیب مستخرج از مدل رقومی ارتفاع و مقادیر انعکاس سطحی باند چهارم استفاده گردیده است.

نحوه کارکرد مدل همان گونه که پیشتر اشاره شد به صورت شاخه ای است که از پنج سطح تصمیم گیری تشکیل یافته است که در نهایت به تعیین شش کلاس اراضی کشاورزی، اراضی جنگلی، اراضی مسکونی شهر یاسوج، آب های سطحی و اراضی بایر و صخره ای می انجامد. در اجرای مدل مفهومی پس از اعمال پیش پردازش ها ابتدا همان گونه که در روندنما مشاهده می شود با استفاده از SAVI و تعیین یک حد آستانه برای آن، پوشش گیاهی از سایر اطلاعات روی تصاویر جدا می شود. علت استفاده SAVI در این قسمت را می توان دخالت عامل خاک دانست. در شاخه صحیح گره تصمیم گیری اول اطلاعات پوشش گیاهی نتیجه کار است که با استفاده از اشتراک دو حد آستانه مشخص شده برای شاخص NDVI و مقادیر انعکاس سطحی باند ۴ سنجنده TM اراضی کشاورزی از اراضی جنگلی مجزا می گردند. در شاخه نادرست گره تصمیم گیری اول اطلاعات بدون پوشش گیاهی موجود است که با به کارگیری اشتراک دو حد آستانه تعیین شده برای شاخص NDWI و مقادیر انعکاس سطحی باند ۴ سنجنده TM آب نیز از اطلاعات موجود در روی تصویر جدا می شود. حاصل شاخه نادرست سطح تصمیم گیری دوم اطلاعاتی است که فاقد آب و پوشش گیاهی است. در سطح تصمیم گیری بعدی با استفاده از اشتراک دو حد مشخص شده برای شاخص NDBI و مقادیر دمای سطح زمین اراضی مسکونی شهر یاسوج هم تعیین کلاس می گردد. در آخرین سطح تصمیم گیری که در واقع اطلاعات شاخه درست سطح تصمیم گیری قبلی یا به عبارتی اطلاعات فاقد پوشش گیاهی، آب و اراضی مسکونی شهر یاسوج را دریافت می کند بدین ترتیب عمل گردید که ابتدا بر روی اطلاعات ورودی این سطح طبقه بندی حداکثر شباهت انجام گرفت و سپس با

اشتراک نتایج این طبقه‌بندی با حد آستانه‌ای که به عنوان حداکثر شیب برای زمین‌های بایر در نظر گرفته شده است اراضی صخره‌ای از اراضی بایر مجزا گردید.

نحوه‌ی تعیین حدود آستانه

برای تعیین حدود آستانه از الگوریتم C5.0 که به طور گسترده‌ای برای امور داده‌کاوی و طبقه‌بندی به کار گرفته می‌شود، استفاده شده است. این الگوریتم را می‌توان شکل تغییر یافته ID3 دانست که اولین بار توسط Quinlan (1986) ارائه گردید. C5.0 الگوریتمی است که به عنوان یک ابزار داده‌کاوی با قابلیت به کارگیری مشخصه‌های پیوسته، برای برقراری قاعده‌های Decision Tree استفاده می‌شود. C5.0 با تعیین شاخه‌ها در مجموعه داده به گونه‌ای که آنتروپی در یک گره کم‌ترین حد ممکن باشد، Decision Tree را ایجاد می‌کند. تصمیم‌گیری در یک گره مشخص به گونه‌ای که بهترین انشعاب حاصل گردد بر مبنای معیار نسبت بهره (gain ratio) صورت می‌گیرد، که عبارتست از نسبت بهره به اطلاعات شاخه (split info). بهره در واقع تغییر در آنتروپی بین گره و آنتروپی حاصل از تمامی زیرگره‌های ریشه گرفته از شاخه است. بهره شاخه X به صورت زیر محاسبه می‌گردد:

$$\text{gain}(X) = \text{info}(T) - \sum_{i=1}^n \frac{|T_i|}{|T|} \text{info}(T_i)$$

که T موارد تعلیمی در گره، T_i موارد تعلیمی در زیرگره i-ام تحت شاخه X است. $\text{info}(T)$ و $\text{info}(T_i)$ به ترتیب میانگین اطلاعات مجموعه‌های T و T_i هستند. رابطه:

$$\text{info}(S) = - \sum_{j=1}^K \frac{\text{freq}(C_j, S)}{|T|} \log_2 \left(\frac{\text{freq}(C_j, S)}{|T|} \right)$$

برای مجموعه S بوده و Cj کلاس j-ام را مشخص می‌کند. اطلاعات شاخه برای شاخه X بوسیله رابطه زیر تعیین می‌گردد:

$$\text{split info}(X) = - \sum_{i=1}^n \frac{|T_i|}{|T|} \log_2 \left(\frac{|T_i|}{|T|} \right)$$

نسبت بهره، نسبت اطلاعات تولید شده بوسیله شاخه‌ای که برای طبقه‌بندی مناسب است را می‌سنجد. شاخه‌ای انتخاب می‌شود که نسبت بهره و در نتیجه اطلاعات بدست آمده را حداکثر سازد. در این تحقیق از الگوریتم C5.0 فقط جهت تعیین حدود آستانه استفاده گردید. برای این منظور نرم افزار See5.0 و داده‌های کنترلی استفاده شد. این داده‌ها با نسبت سه به یک به داده‌های تعلیمی و تست تقسیم شدند. سپس با استفاده از داده‌های تعلیمی حدود آستانه تعیین شده و به کمک داده‌های تست خطای آن تخمین زده شد که بیشترین مقدار آن برای عامل شیب ۱/۷ درصد بود. شکل ۴ حد آستانه محاسبه شده SAVI برای داده‌های مربوط به سال ۱۹۹۰ را نشان می‌دهد.

```
See5 [Release 2.02] Fri Jan 01 20:43:00 2010
Class specified by attribute `Groups'
Read 846 cases (2 attributes) from SAVI_Threshold.data
Decision tree:
SAVI <= 0.30100058: Non-Vegetation_Data (553)
SAVI > 0.30100058: Vegetation_Data (293)
```

شکل ۴- نحوه تعیین حد آستانه SAVI برای داده‌های سال ۱۹۹۰

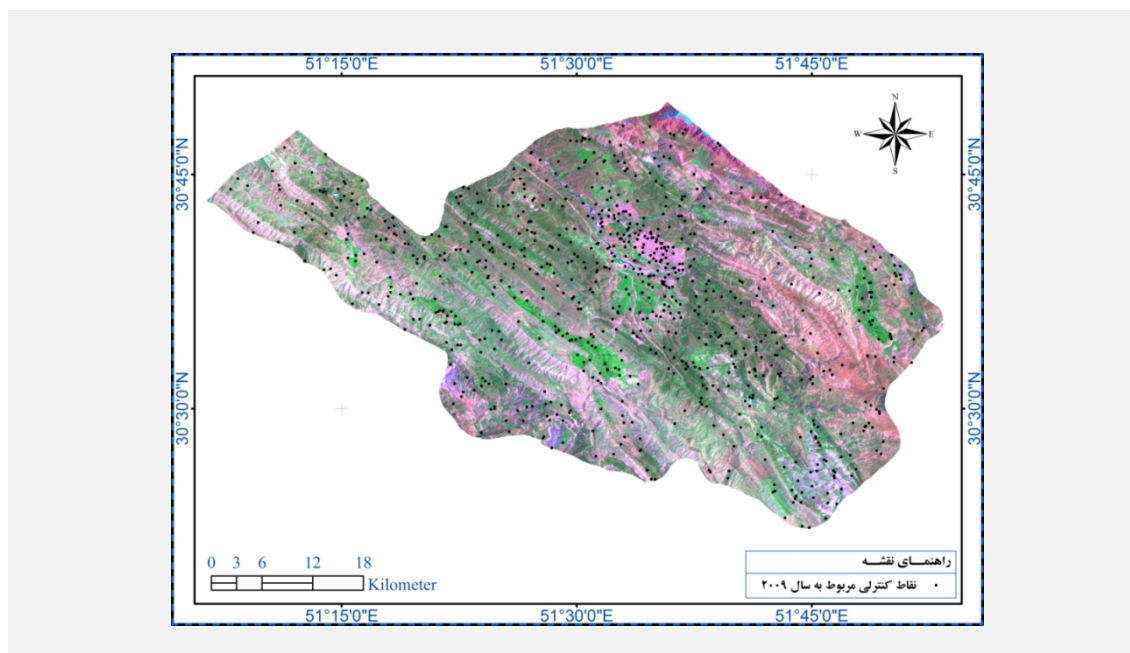
ارزیابی صحت تصاویر طبقه‌بندی شده

ارزیابی نتایج طبقه‌بندی یکی از مراحل مهم پس از طبقه‌بندی^{۱۴} است. ارائه نتایج طبقه‌بندی بدون هیچ گونه پارامتری که کیفیت و صحت این نتایج را بیان کند از ارزش آن‌ها می‌کاهد (Fatemi and Rezaii, 2005). برای تعیین صحت تصویر طبقه‌بندی شده ۱۹۹۰ از نقاط کنترلی برداشت شده از تصاویر ماهواره‌ای اسپات همان سال استفاده گردید و هم‌چنین برای ارزیابی صحت تصویر طبقه‌بندی شده ۲۰۰۹ از نقاط واقعیت زمینی در منطقه که توسط GPS برداشت شده‌اند، استفاده شد. شکل ۵ نحوه پراکنش نقاط GPS را در تصویر سال ۲۰۰۹ را نشان می‌دهد. ماتریسهای خطا بصورت جدول‌بندی متقاطع از کلاس‌های طبقه‌بندی شده در برابر داده‌های مرجع برای ارزیابی صحت طبقه‌بندی بکار گرفته شدند (Congalton and Green, 1999). صحت کلی، صحت‌های کاربر و تولید کننده و آماره کاپا از ماتریس‌های خطا استخراج شدند.

مقدار آماره کاپا صحت طبقه‌بندی را نسبت به یک طبقه‌بندی کاملاً تصادفی محاسبه می‌نماید (Richrds and Jia, 2006) و مقدار آن از فرمول زیر بدست می‌آید.
(توافق مورد انتظار - ۱) ÷ (توافق مورد انتظار - توافق مشاهده شده) = مقدار آماره کاپا

آشکارسازی تغییرات

در این تحقیق با استفاده از تصاویر حاصل از طبقه‌بندی و به کمک الگوریتم آشکارسازی تغییرات به روش مقایسه‌ی پس از طبقه‌بندی^{۱۵} تغییرات پوشش اراضی برای دوره بیست ساله ۱۹۹۰ تا ۲۰۰۹ مشخص گردید. این تکنیک، رایج‌ترین روش برای آشکارسازی تغییرات است (Jensen, 2004). در این روش، اطلاعات مربوط به مقدار تغییرات "از - به" یعنی از نوع مشخصی از پوشش اراضی به نوع دیگری از آن و هم‌چنین نوع تغییرات چشم انداز به آسانی می‌تواند محاسبه شده و به صورت نقشه هم نمایش داده شود.



شکل ۵- پراکنش نقاط برداشت شده با GPS در منطقه مطالعاتی

ارزیابی صحت آشکارسازی تغییرات

در مورد ارزیابی صحت آشکارسازی تغییرات مشکلات عدیده‌ای وجود دارد چون نمونه برداری از مناطقی که در آینده تغییر خواهند کرد، قبل از آنکه این نواحی تغییر کنند کار دشواری است (Congalton and Green, 1999). یک موضوع مهم در تحلیل آشکارسازی تغییرات خطاهای مربوط به موقعیت و خصوصیت است که در تصاویر چند زمانه اهمیت می‌یابد مخصوصاً هنگامی که بیش از دو تاریخ در تحلیل به کار گرفته شود. ساده‌ترین روش برای ارزیابی صحت نقشه‌های تغییرات، ضرب نمودن صحت‌های مربوط به تصاویر طبقه‌بندی شده‌ی دخالت داده شده در آشکارسازی تغییرات است.

روش اشاره شده علی‌رغم سرعت و سهولت، دقت بالایی ندارد. یک روش دقیق برای ارزیابی صحت آشکارسازی تغییرات، نمونه‌برداری تصادفی از مناطق طبقه‌بندی شده به عنوان "تغییر کرده" و "تغییر نکرده" و سپس مشخص نمودن اینکه آیا آن‌ها به درستی طبقه‌بندی شده‌اند، است (Fuller et al., 2003). ما در این پژوهش با توجه به دقت بالاتر روش دوم، علاوه بر روش اول، آن‌را نیز برای ارزیابی صحت آشکارسازی تغییرات بکار گرفته‌ایم. بدین منظور حجم نمونه (N) از فرمول زیر تعیین گردید.

$$N = Z^2 \times P \times (1-P) / E^2$$

که Z مقدار ارزش استاندارد (به عنوان مثال ۱/۹۶ برای سطح اطمینان ۹۵ درصد)، P دقت مورد انتظار و E خطای مجاز است. برای دقت مورد انتظار ۵۰ درصد، سطح اطمینان ۹۵ درصد و خطای مجاز ۵ درصد حجم نمونه به تعداد ۳۸۴ از فرمول فوق محاسبه شد. از این رو تعداد ۳۸۴ پیکسل از هر کلاس "تغییر کرده" و "تغییر نکرده" به صورت تصادفی انتخاب گردید. از آنجایی که نمونه برداری به صورت تصادفی انجام گرفته بود

تعدادی از پیکسل‌های انتخابی در مناطق مرزی دو کلاس مذکور قرار گرفته بودند که برای افزایش دقت این نمونه‌ها حذف شدند و بدین ترتیب ۳۳۱ نمونه برای کلاس "تغییر کرده" و ۳۶۳ نمونه برای کلاس "تغییر نکرده" باقی ماند. هر نقطه نمونه‌برداری شده برای تعیین اینکه آیا تغییرات صورت گرفته است یا نه، با داده‌های مرجع SPOT برای سال ۱۹۹۰ و برداشتهای سیستم موقعیت‌یابی جهانی برای سال ۲۰۰۹ مقایسه گردید. این روش بسته به دوره‌هایی که برای آشکارسازی تغییرات به کار گرفته شده است به تفسیر بصری و حجم کار بالایی نیاز دارد، ولی در کنار این امر دقت بالا از امتیازات مهم آن می‌باشد.

نتایج

صحت طبقه‌بندی و صحت آشکارسازی تغییرات

در این تحقیق ماتریس خطا برای ارزیابی صحت طبقه‌بندی تصاویر سال‌های ۱۹۹۰ و ۲۰۰۹ استفاده گردید که نتایج در جداول ۲ و ۳ نمایش داده شده‌اند. صحت کل برای سال‌های ۱۹۹۰ و ۲۰۰۹ به ترتیب ۹۳/۱۶ درصد و ۹۳/۳۷ درصد و مقدار آماره کاپا ۹۳ درصد و ۹۱/۸۷ درصد است. همانگونه که در جداول مذکور مشاهده می‌گردد صحت کاربر و تولیدکننده برای طبقه‌بندی هر دو سال بالا بوده و بین ۸۲ تا ۹۷ درصد قرار می‌گیرد. در مقایسه با تحقیقات مشابهی که با استفاده از روش طبقه‌بندی نظارت شده و به کمک الگوریتم‌های معمول مانند حداکثر مشابهت و داده‌های TM برای آشکارسازی تغییرات صورت گرفته است، بهره‌گیری از مدل مفهومی مبتنی بر الگوریتم Decision Tree تا حد زیادی صحت کلی، صحت‌های کاربر و تولیدکننده و آماره کاپا را افزایش داده است که نشان از کارایی بالایی این مدل در امر طبقه‌بندی دارد.

جدول ۲- خلاصه‌ای از صحت طبقه‌بندی تصویر TM لندست برای سال ۱۹۹۰

کلاس‌ها	اراضی بایر	شهر	اراضی کشاورزی	اراضی جنگلی	آب	اراضی صخره‌ای	کل	صحت کاربر
اراضی بایر	۲۰۲	۲	۳	۱	۱	۵	۲۱۴	۹۴/۳۹
شهر	۸	۹۲	۳	۲	۰	۷	۱۱۲	۸۲/۱۴
اراضی کشاورزی	۲	۰	۱۰۲	۴	۰	۰	۱۰۸	۹۴/۴۴
اراضی جنگلی	۴	۱	۱۲	۲۶۳	۱	۱	۲۸۲	۹۳/۲۶
آب	۱	۰	۱	۰	۴۴	۰	۴۶	۹۵/۵۶
اراضی صخره‌ای	۱۱	۱	۲	۳	۱	۳۴۷	۳۶۵	۹۵/۰۶
کل	۲۲۸	۹۶	۱۲۳	۲۷۳	۴۷	۳۶۰	۱۱۲۷	-
صحت تولیدکننده	۸۸/۵۹	۹۵/۸۳	۸۲/۹۲	۹۶/۳۳	۹۳/۶۱	۹۶/۳۸	-	-

جدول ۳- خلاصه‌ای از صحت طبقه‌بندی تصویر TM لندست برای سال ۲۰۰۹

کلاس‌ها	اراضی بایر	شهر	اراضی کشاورزی	اراضی جنگلی	آب	اراضی صخره‌ای	کل	صحت کاربر
اراضی بایر	۱۴۴	۲	۲	۱	۱	۴	۱۵۴	۹۳/۵
شهر	۵	۷۰	۲	۰	۰	۳	۸۰	۸۷/۵
اراضی کشاورزی	۲	۰	۷۴	۲	۰	۰	۷۸	۹۴/۸۷
اراضی جنگلی	۳	۱	۷	۱۸۹	۱	۲	۲۰۳	۹۳/۱
آب	۱	۰	۰	۰	۳۷	۰	۳۸	۹۷/۳۶
اراضی صخره‌ای	۱۲	۱	۱	۱	۰	۲۴۷	۲۶۲	۹۴/۲۷
کل	۱۶۷	۷۴	۸۶	۱۹۳	۳۹	۲۵۶	۸۱۵	-
صحت تولیدکننده	۸۶/۲۲	۹۴/۵۹	۸۶/۰۴	۹۷/۹۲	۹۴/۸۷	۹۶/۴۸	-	-

تغییرات استخراج شد (جدول ۴). صحت کلی برای آشکارسازی تغییرات ۸۵/۶ درصد و مقدار آماره کاپا ۷۱/۱ درصد تعیین گردید.

۳-۲- نقشه‌ها و آمار مربوط به طبقه‌بندی و آشکارسازی تغییرات

نقشه‌های طبقه‌بندی برای هر دو سال تولید شدند (شکل ۶ و ۷) که مساحت‌های هر کدام از کلاس‌ها و مقدار تغییرات آن‌ها برای دوره بیست ساله در جدول ۵ خلاصه شده است.

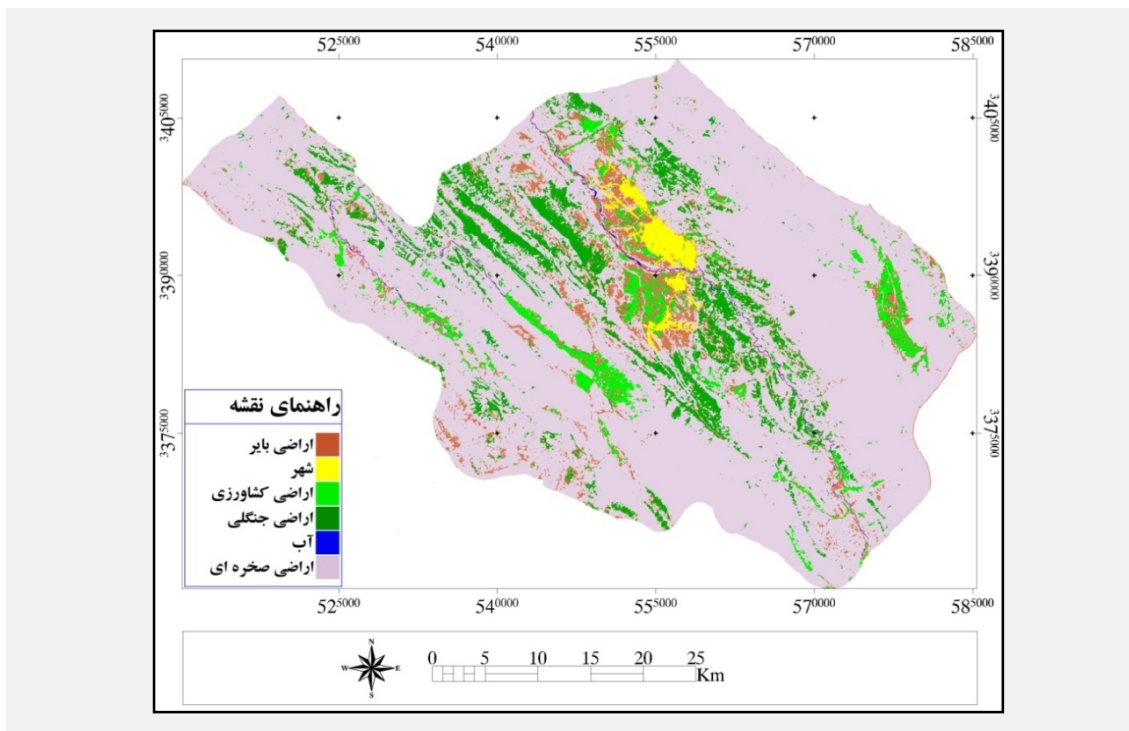
برای ارزیابی صحت آشکارسازی تغییرات هر دو روشی که در قسمت ۲-۳-۶ توضیح داده شدند به کار گرفته شد. نتیجه ضرب صحت‌های مربوط به تصاویر طبقه‌بندی شده‌ی دخالت داده شده در آشکارسازی تغییرات به عنوان حاصل روش اول ۸۶/۹۸ درصد به دست آمد. همچنین صحت آشکارسازی تغییرات با روش دوم که در آن نمونه تصادفی به عنوان "تغییر کرده" یا "تغییر نکرده" در نقشه تغییرات در نظر گرفته شدند ارزیابی گردید و ماتریس خطای آشکارسازی

جدول ۴- ماتریس خطای آشکارسازی تغییرات

صحت کاربر (به درصد)	طبقه بندی		کلاسهای مرجع
	تغییر نکرده	تغییر کرده	
۸۴/۲۹	۵۲	۲۷۹	تغییر کرده
۸۶/۷۷	۳۱۵	۴۸	تغییر نکرده
	۸۵/۸۳	۸۵/۳۲	صحت تولید کننده (به درصد)
		۸۵/۶	صحت کلی (درصد)
		۷۱/۱	آماره کاپا (درصد)

جدول ۵- مساحت‌های کلاس‌ها و مقدار تغییرات در دو دهه

تغییرات از ۱۹۹۰ تا ۲۰۰۹		سال ۲۰۰۹		سال ۱۹۹۰		کلاس پوشش اراضی	
درصد تغییرات نسبی	درصد از کل	مساحت (کیلومتر مربع)	مساحت (کیلومتر مربع) (درصد)	مساحت (کیلومتر مربع) (درصد)	مساحت (کیلومتر مربع)		
۲۱۹/۱۹	۰/۹۶	۱۹/۲۸	۱/۴	۲۸/۰۷	۰/۴۴	۸/۷۹	اراضی مسکونی
-۳۱/۷۶	-۴/۷۹	-۹۵/۷۳	۱۰/۲۸	۲۰۵/۷۱	۱۵/۰۶	۳۰۱/۴۵	اراضی جنگلی
-۷/۳۲	-۰/۳۱	-۶/۲۴	۳/۹۵	۷۹/۱	۴/۲۶	۸۵/۳۴	اراضی کشاورزی
۵۱/۵۵	۲/۶۴	۵۲/۸	۷/۷۵	۱۵۵/۲۲	۵/۱۲	۱۰۲/۴۲	اراضی بایر
۴/۶۷	۰/۱	۰/۲۵	۰/۳۶	۵/۵۵	۰/۲۶	۵/۳	آب
۲	۱/۴۹	۳۰	۷۶/۳۴	۱۵۲۸/۰۳	۷۴/۵۸	۱۴۹۸/۰۳	اراضی صخره‌ای
-	-	-	۱۰۰	۲۰۰۱/۳۳	۱۰۰	۲۰۰۱/۳۳	جمع



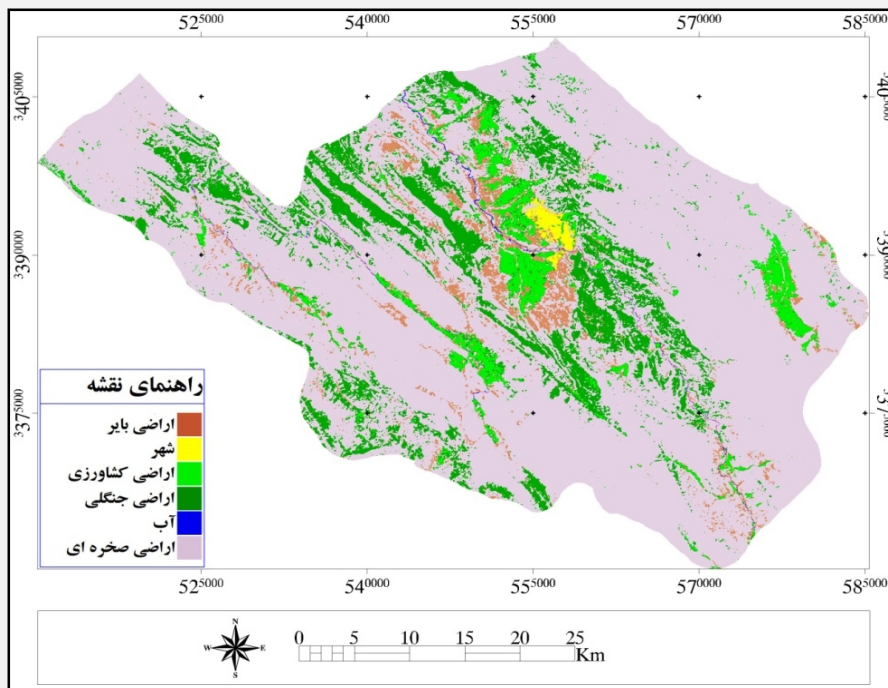
شکل ۶ - تصویر طبقه بندی شده سال ۱۹۹۰ با استفاده مدل مفهومی مبتنی بر الگوریتم Decision Tree

با توجه به اینکه نقشه تغییرات در مقیاس کوچک برای منطقه مورد مطالعه نمی تواند به خوبی تغییرات ایجاد شده را در این دو دهه نشان دهد از این رو تنها به نمایش نقشه تغییرات از کلاس های مختلف به کلاس اراضی مسکونی در محدوده شهر یاسوج در شکل ۸ اکتفا می کنیم.

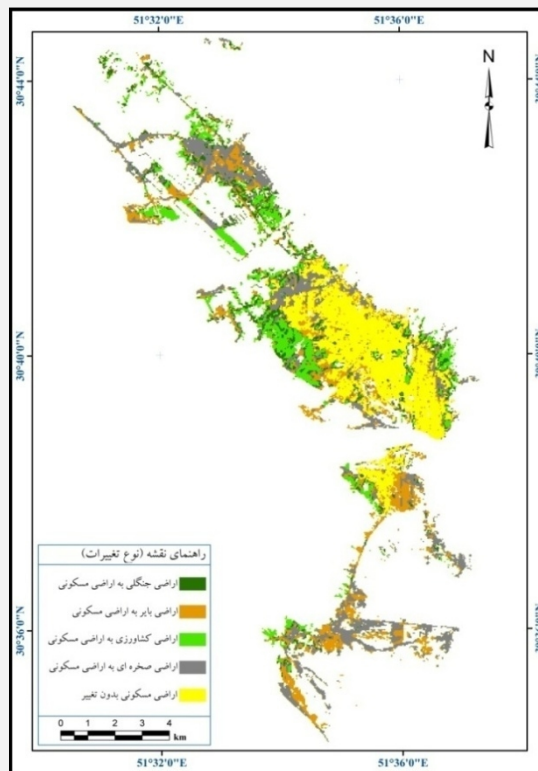
تحلیل الگوهای تغییر

اگرچه تا حدی می توان به آمارهای مشابه آنچه که بیان شد از طریق شهرداری ها یا مراکز آمار و ... دست یافت ولی تنها پردازش تصاویر سنجش از دوری است که با هزینه اندک، زمان کم و دقت قابل قبول این اطلاعات را مهیا کرده و پاسخ این سوال که "پوشش اراضی یا کاربری اراضی در کجا تغییر کرده است؟" را می دهد.

از سال ۱۹۹۰ تا ۲۰۰۹ مساحت منطقه مسکونی حدود ۱۹/۲۸ کیلومترمربع (۰/۹۶ درصد) افزایش یافته، در حالی که اراضی کشاورزی و جنگلی به ترتیب ۶/۲۴ کیلومترمربع (۰/۳۱ درصد) و ۹۵/۷۳ (۴/۷۹ درصد) کاهش مساحت داشته اند. اراضی بایر رشدی معادل ۵۲/۸ کیلومترمربع (۲/۶۴ درصد) داشته است. بصورت نسبی نیز اراضی مسکونی در این دو دهه ۲۱۹/۱۹ درصد افزایش داشته اند (۸/۷۹ کیلومترمربع در سال ۱۹۹۰ و ۲۸/۰۷ کیلومترمربع در سال ۲۰۰۹). اراضی بایر نیز افزایش ۵۱/۵۵ درصدی داشته اند، در حالی که اراضی کشاورزی و جنگلی به ترتیب یک کاهش ۷/۳۲ و ۳۱/۷۶ درصدی را نشان می دهند. در مقایسه با سایر کلاس ها دو کلاس اراضی صخره ای و آب بصورت نسبی تغییرات بسیار اندکی را به نمایش گذاشته اند.



شکل ۷- تصویر طبقه‌بندی شده سال ۲۰۰۹ با استفاده مدل مفهومی مبتنی بر الگوریتم Decision Tree



شکل ۸- نقشه تغییرات از کلاس‌های مختلف به کلاس اراضی مسکونی در محدوده شهر یاسوج
 علوم محیطی سال هشتم، شماره سوم، بهار ۱۳۹۰
 ENVIRONMENTAL SCIENCES Vol.8, No.3, Spring 2011

از این رو با تولید یک نقشه تغییرات مشابه آنچه که در شکل ۸ مشاهده می شود، مزایای سنجش از دور ماهواره ای در جداسازی مکانی آمار تغییرات می تواند به خوبی نمایان گردد. شکل ۸ تغییر از سه پوشش اراضی بایر، اراضی کشاورزی و اراضی جنگلی را به منطقه مسکونی نشان می دهد. همان گونه که در این شکل دیده می شود منطقه شهری در جهت شمال غربی و جنوب رشد داشته است. رشد شهر و کاهش اراضی جنگلی و کشاورزی از مهم ترین تغییرات در ناحیه نشان داده شده است. اگرچه شکل ۸ فقط تغییرات بیان شده را نشان می دهد ولی به صورت مشابهی می توان سایر تغییرات را هم نمایش داد.

بحث

از آنجا که در اغلب مطالعات برای تهیه نقشه کاربری اراضی از عکس های هوایی و تصاویر ماهواره ای استفاده می شود، پژوهش حاضر نیز با هدف توسعه یک مدل مفهومی مبتنی بر الگوریتم DT و داده کاوی برای آشکارسازی تغییرات پوشش و کاربری اراضی در بخش مرکزی شهرستان بویراحمد برای یک دوره زمانی ۲۰ ساله انجام شد. برای انجام تحقیق از تصاویر ماهواره ای چند زمانه سنجنده TM ماهواره لندست و داده های کمکی استفاده شده و پس از تهیه نقشه کاربری اراضی هر دوره زمانی، نقشه تغییرات پوشش و کاربری اراضی استخراج گردید. نتایج این تحقیق نشان می دهد که داده های سنجش از دور همراه با تکنیک های داده کاوی از توانایی بالایی در استخراج انواع نقشه های کاربری اراضی و هم چنین ارزیابی تغییرات کاربری برخوردارند و هم چنین پتانسیل داده های سنجنده TM ماهواره لندست را به عنوان یک ابزار مناسب و اقتصادی برای تصویر نمودن و تجزیه و تحلیل تغییرات پوشش اراضی در طی زمان نشان می دهد که از این داده ها یا داده های جانشین

آن یعنی ETM+ و Advanced Land Imager یا ALI می توان به عنوان ورودی هایی برای مدیریت سرزمین و سیاست گذاری ها استفاده کرد. هم چنین نتایج پژوهش حاضر حاکی از آن است که استفاده از روش شاخه ای یا چند مرحله ای برای طبقه بندی تصاویر ماهواره ای، دارای مزایایی از جمله کاستن زمان پردازش، بهبود صحت کلاس های تعلیم کوچک، بکارگیری منابع متفاوتی از داده ها، مجموعه های مختلف از خصوصیات و حتی الگوریتم های متفاوت در هر مرحله از تصمیم گیری است.

تشکر و قدردانی

نویسندگان مقاله مراتب سپاس و قدردانی خود را از پروفیسور Ross Quinlan به خاطر راهنمایی های ارزنده ایشان و در اختیار قرار دادن نرم افزار See5.0 اعلام می دارند.

پی نوشت ها

1. Bidirectional Reflectance Distribution Function (BRDF)
2. Illumination
3. Maximum likelihood
4. Minimum distance Classifier
5. Surface reflectance
6. Flowchart
7. Normalized Difference Vegetation Index
8. Soil Adjusted Vegetation Index
9. Texture
10. Entropy
11. Normalized Difference Building-up Index
12. Land Surface Temperature (LST)
13. Ground surface emissivity
14. Post classification
15. Post-classification comparison change detection algorithm
16. From-to

- Feyzi zade, B. and S.M. Haj mir rahimi (2007). Change detection of green spaces of Tabriz city by using object oriented method, Urban GIS Conference.
- Fuller, R.M., G.M. Smith and B.J. Devereux (2003). The characterization and measurement of land cover change through remote sensing: Problems in operational applications. *International Journal of Applied Earth Observation and Geoinformation*, 4: 243–253.
- Goetz, S.J., D. Varlyguin, A.J. Smith, R.K. Wright, S.D. Prince and M.E. Mazzacato (2004). Application of multitemporal Landsat data to map and monitor land cover and land use change in the Chesapeake Bay watershed, *Proceedings of the second international workshop on the analysis of multi-temporal remote sensing images*, 223– 232.
- Hansen, M., R. Dubayah, and R. Defries (1996). Classification trees: an alternative to traditional land cover classifiers. *International Journal of Remote Sensing*, 17: 1075-1081.
- Huete, A.R. (1988). A soil-Adjusted Vegetation Index (SAVI). *Remote Sensing of Environment*, 25: 295-309.
- Jensen, J.R. (2004). *Digital change detection: Introductory digital image processing: A remote sensing perspective*. New Jersey: Prentice-Hall.
- Alavi panah, S.K. (2006). *Thermal remote sensing & its application in earth science*. publication of university of Tehran, Tehran.
- Alavi panah, S.K. and A.H. Ehsani (2005). Change detection and desertification based on geographic information system and multi-spectral satellite data (Case study: damghan playa). *Geographical research quarterly*, 51: 143-155.
- Alberti, M., R. Weeks and S. Coe (2004). Urban land cover change analysis in Central Puget Sound. *Photogrammetric Engineering and Remote Sensing*, 70(9): 1043– 1052.
- Bahadur, R. and Y. Murayama (2006). Land use change analysis using remote sensing and GIS: A case study of Kathmandu metropolitan, Nepal. *Research Abstracts on Spatial Information Science*, 22: 167-179.
- Brodley, C.E., T. Lane and T.M. Stough (1999). *Knowledge Discovery and Data Mining*. *American Scientist*, 87: 54-61.
- Congalton, R.G. and K. Green (1999). *Assessing the accuracy of remotely sensed data: Principles and practices*. Florida: Lewis Publishers.
- Fatemi, S.B. and Y. Rezaii (2005). *Principles of remote sensing*. Tehran: Azade publication.

- Richards, J.A. and X. Jia (2006). Remote Sensing Digital Image Analysis: An Introduction. Berlin: Springer-Verlag.
- Rounds, E.M. (1980). A Combined Nonparametric Approach to Feature Selection and Binary Decision Tree Design, *Pattern Recognition*, 12: 313–317.
- Seto, K.C., C.E. Woodcock, c. song and R.k. kaufmann (2002). Monitoring land-use change in the Pearl River Delta using Landsat TM. *International Journal of Remote Sensing*, 8(6): 460–471.
- Swain P.H. and H. Hauska (1977). The Decision Tree Classifier: Design and Potential. *Geoscience Electronics*, GE-15: 142–147.
- Vogelmann, J.E. (2001). Completion of the 1990s national land cover data set for the conterminous United States from Landsat Thematic Mapper and ancillary data sources. *Photogrammetric Engineering & Remote Sensing*, 67: 650-662.
- Yang, X. (2002). Satellite monitoring of urban spatial growth in the Atlanta metropolitan area, *Photogrammetric Engineering and Remote Sensing*, 68 (7): 725–734.
- Yong, Z., N.I. Shao-xiang and Y. Shan (2003). An Effective Approach to Automatically Extract Urban Land-use from TM Imagery. *Journal of Remote Sensing*, 7(1): 37–40.
- Jiang, Q. and H. Liu (2004). Extracting TM Image Information Using Texture Analysis. *Journal of Remote Sensing*, 8(5): 459- 464.
- Jimenez-Munoz, J.C. and J.A. Sobrino (2003). A generalized single- channel method for retrieving land surface temperature from remote sensing data. *Journal of Geophysical Research*, 108: 4688-4698.
- Kaya, S. and P.J. Curran (2006). Monitoring urban growth on the European side of the Istanbul metropolitan area. *International Journal of Applied Earth Observation and Geoinformation*, 8: 18–25.
- Kemper, G., O. Altan., P.t. kov and C. Lavallo (2001). Monitoring Landuse change of the city of greater Istanbul. (Forth international Symposium “Turkish-German Joint Geodetic Days, Vol 2: Berlin.
- Lawrence, R.L. and A. Wright (2001). Rule-Based Classification systems using classification and regression tree (Cart) analysis. *Photogrammetric Engineering & Remote Sensing*, 67:1137-1142.
- McFeeters, S.K. (1996). The use of the normalized difference water index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17: 1425–1432.
- Mui, J.K. and K.S. Fu (1980). Automated Classification of Nucleated Blood Cells Using a Binary Tree Classifier. *Pattern Analysis and Machine Intelligence*, 2: 429–443.
- Quinlan, J.R. (1986). Induction of decision trees. *Machine Learning*, 1: 81–106.



