

پیش‌بینی مکانی عملکرد گندم با استفاده از نقشه‌برداری رقومی خاک در منطقه گنوند استان خوزستان

روح‌اله تقی زاده مهرجردی^{۱*}، سیدعلیرضا سیدجلالی^۲، فریدون سرمدیان^۳

۱. استادیار دانشکده کشاورزی و منابع طبیعی دانشگاه اردکان

۲. استادیار مؤسسه تحقیقات خاک و آب

۳. استاد گروه مهندسی علوم خاک، دانشکده مهندسی و فناوری کشاورزی، پردیس کشاورزی و منابع طبیعی

(تاریخ دریافت: ۱۳۹۳/۱۲/۱۷ - تاریخ تصویب ۱۳۹۴/۶/۲۳)

چکیده

در پژوهش حاضر، مقدار ۱۱۰ عملکرد مشاهداتی در منطقه گنوند (استان خوزستان) با استفاده از مدل برنامه‌ریزی ژنتیک، به داده‌های کمکی (مستخرج شده از مدل رقومی ارتفاع و تصویر ماهواره) ارتباط داده شد. سپس با استفاده از معادله بدست آمده برای نقاط فاقد مشاهده میزان عملکرد برآورد و نقشه تغییرات مکانی محاسبه گردید. الگوریتم رپر پارامترهای مؤلفه تصویر ماهواره، شاخص نسبت گیاهی، شاخص گیاهی تعدیل‌کننده اثر خاک، شاخص خیزی و سطح پایه شبکه زهکشی را به عنوان مهم‌ترین عوامل تولید شناخته است. میانگین ریشه مربعات خطا، ضریب همبستگی تطابق لاین و ضریب تبیین ارزیابی متقابل مدل برنامه‌ریزی ژنتیک (۱) با همه داده‌های کمکی به ترتیب ۵۲۵/۱۱، ۰/۸۷ و ۰/۸۲ است. همچنین نتایج نشان داد که مدل برنامه‌ریزی ژنتیک (۲) با داده‌های کمکی انتخاب شده توسط الگوریتم رپر نیز به خوبی (میانگین ریشه مربعات خطا، ضریب همبستگی تطابق لاین و ضریب تبیین به ترتیب ۵۳۰/۸۲، ۰/۸۶ و ۰/۷۹ است) قادر به پیش‌بینی عملکرد گندم است. لذا پیشنهاد می‌گردد در مطالعات آتی جهت برآورد مکانی عملکرد محصولات زراعی از مدل برنامه‌ریزی ژنتیک در قالب نقشه‌برداری رقومی خاک استفاده گردد.

واژه‌های کلیدی: داده‌های کمکی، الگوریتم رپر، برنامه‌ریزی ژنتیک، تغییرات مکانی

مقدمه

گندم یکی از غلات کلیدی است که در سراسر جهان کشت می‌شود و کالری اولیه و منبع تغذیه برای میلیون‌ها نفر از مردم جهان را فراهم می‌کند (Becker-Reshef *et al*, 2010). مطالعات اخیر نشان می‌دهد که تقاضای جهانی برای مواد غذایی به مدت حداقل تا ۴۰ سال دیگر رو به افزایش است. جمعیت جهان در اواسط این قرن به ۹ میلیارد نفر خواهد رسید (Charles *et al*, 2010). افزایش تقاضا برای مواد غذایی در سراسر جهان و منابع محدود در دسترس برای تولید، لزوم استفاده از ابزارهای جدید برای برآورد تولید محصول را می‌طلبد (Padilla *et al*, 2012). میانگین عملکرد واقعی محصولات زراعی بطور سالیانه برای مناطق مختلف کشور توسط وزارت جهاد کشاورزی اندازه‌گیری می‌شود و هیچگونه اطلاعاتی در خصوص پایین بودن عملکرد در واحد سطح وجود ندارد. و در ضمن کشاورزان و برنامه‌ریزان هیچگونه اطلاعاتی از محدودیت‌هایی که باعث پایین بودن عملکرد در واحد سطح اراضی آن‌ها می‌شوند را در اختیار ندارند.

استفاده از مدل‌های پیش‌بینی عملکرد محصول می‌تواند عملکرد محصول را با توجه به محدودیت‌های مختلف محیطی از قبیل اقلیم، خاک، زمین‌نما^۱ و مدیریت تعیین کند و راه‌گشایی برای حل این مشکل باشد.

برای پیش‌بینی عملکرد محصول قبل از برداشت، طیف وسیعی از روش‌ها مانند برآوردهای چشمی، مدل‌های شبیه‌سازی محصول و روش‌های رگرسیون وجود دارد (Chipanshi *et al*, 1999; Doraiswamy *et al*, 2003; Maselli *et al*, 2001; Wall *et al*, 2007). مدل‌های محصول غالباً در اکولوژی، زراعت و علوم محیطی برای شبیه‌سازی محصول و متغیرهای محیطی استفاده می‌شود. در سال‌های گذشته، مدل‌های شبیه‌سازی رشد محصول به طور گسترده‌ای به عنوان یک ابزار مهم برای بررسی رشد محصول و ارقام مختلف در شرایط مختلف اقلیمی مورد استفاده قرار گرفته است (Toscano *et al*, 2012; Cahn *et al*, 1994). به عنوان مثال، Chen *et al* (2008) از یک مدل رگرسیون خطی برای پیش‌بینی عملکرد گندم زمستانه در استان شاندونگ استفاده کردند (Shahbazi *et al*, 2009) از

بر اساس نظریه تکاملی داروین استوار است. برنامه‌ریزی ژنتیک از روش‌های جدید بین روش‌های الگوریتم‌گردشی است که به دلیل دارا بودن دقت کافی، به عنوان یک روش کاربردی مطرح می‌شود. از برنامه‌ریزی ژنتیک به طور وسیعی در پیش‌بینی ظرفیت تبادل کاتیونی خاک (Taghizadeh-Mehrjardi, 2015)، خصوصیات منحنی رطوبتی خاک (Johari et al, 2006) و بافت خاک (Padarian et al, 2012) استفاده شده است.

با توجه به موارد ذکر شده و اقلیم خشک و نیمه خشک ایران و کمبود عملکرد در واحد سطح و جمعیت رو به افزایش کشور نیاز به تحقیقات جدید و دستیابی به راهکارهایی برای افزایش عملکرد محصول است. بر اساس آمارنامه وزارت جهاد کشاورزی مربوط به سال‌های زراعی ۶۱ لغایت ۸۹ استان خوزستان دارای میانگین عملکرد گندم آبی ۲۵۰۲ کیلوگرم در هکتار بوده که بیشترین عملکرد آن در سال زراعی ۸۳-۸۴ و برابر ۳۵۵۶ کیلوگرم در هکتار و برای گندم دیم ۵۸۵ کیلوگرم در هکتار بوده که حداکثر عملکرد آن در سال زراعی ۸۰-۸۱ برابر ۱۳۰۶ کیلوگرم در هکتار است. برای بالابردن عملکرد گندم در واحد سطح، شاید اولین گام بررسی تغییرات مکانی آن است. بنابراین در تحقیق حاضر با استفاده از داده‌های میدانی، داده‌های کمی و مدل برنامه‌ریزی ژنتیک اقدام به پهنه‌بندی رقوم عملکرد گندم در منطقه گتوند استان خوزستان گردیده است.

مواد و روش‌ها

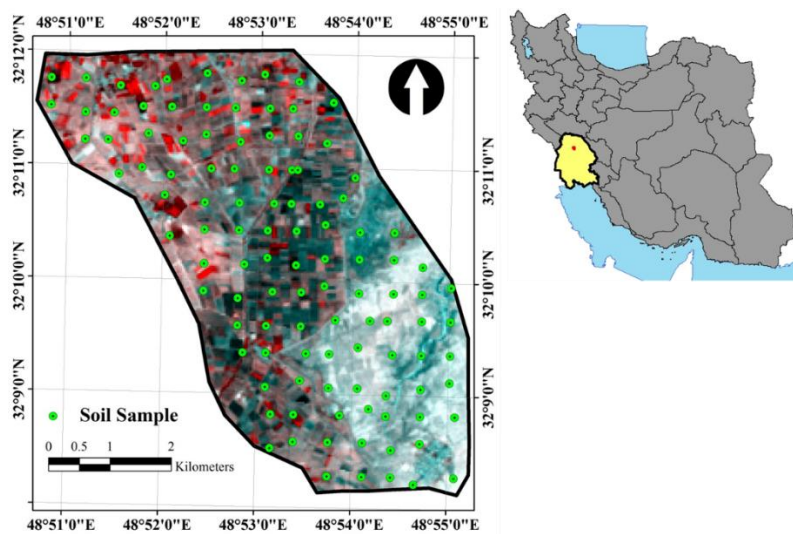
منطقه مورد مطالعه و نمونه‌برداری

منطقه مورد مطالعه در جنوب غرب ایران، دشت عقیلی، گتوند، استان خوزستان به مساحت تقریبی ۳۵۰۰ هکتار بین عرض ۳۲ درجه و هفت دقیقه تا ۳۲ درجه و ۱۰ دقیقه شمالی و طول ۴۸ درجه و ۴۸ دقیقه تا ۴۸ درجه و ۵۶ دقیقه شرقی قرار دارد. با توجه به ایستگاه سینوپتیک شوشتر که نزدیک‌ترین ایستگاه به منطقه مورد مطالعه است، این منطقه بر اساس طبقه‌بندی گوسن^۲ جزء اقلیم نیمه بیابانی خفیف است. حداکثر دمای روزانه در تیر ماه ۴۶/۶ درجه سانتی‌گراد و حداقل دمای روزانه ۸/۱ درجه سانتی‌گراد در دی ماه است. مقدار متوسط بارندگی سالیانه در منطقه مورد مطالعه ۳۲۴ میلی‌متر است. خاک‌های منطقه مورد مطالعه بر اساس روش طبقه‌بندی خاک آمریکا، گروه نقشه‌برداری خاک در دو رده انسپتی‌سول‌ها و ورتیسول‌ها و چهار زیر گروه، Typic Calciustepts, Typic Haplustepts

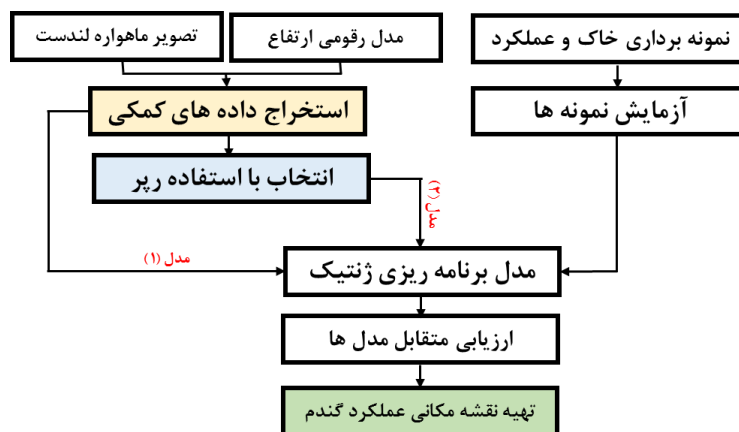
خصوصیات خاک و پارامترهای پستی و بلندی (اولیه و ثانویه) جهت برآورد گندم دیم در منطقه سیلاب استان خراسان بهره گرفته‌اند و بیان داشتند که رابطه رگرسیونی تنها ۴۳ درصد تغییرات عملکرد کل را توجیه می‌کند. دلیل پایین بودن ضریب تبیین احتمالا مربوط به رابطه‌های غیرخطی و پیچیده بین عملکرد و پارامترهای کمکی دارد. لذا جهت بالابردن دقت مدل، لزوم استفاده از مدل‌های پیچیده‌تر احساس می‌شود. در سال-های اخیر محققین در علوم مختلف از انواع روش‌های داده‌کاوی جهت مدل‌سازی و یافتن رابطه‌های پیچیده استفاده کرده‌اند. حتی می‌توان از چنین روابط بدست آمده، برای برآورد مکانی عملکرد گندم بر اساس چارچوب کاری اسکورپن^۱ استفاده نمود. مدل اسکورپن که بر اساس نقشه‌برداری رقوم خاک است به داده‌های نمونه‌برداری شده و مدل‌های استنباط مکانی خاک که می‌توانند برای تخمین و پیش‌بینی متغیر مورد نظر در نقاط نمونه‌برداری نشده به کار روند، توجه دارد. در این مدل هفت عامل شامل ویژگی‌های خاک محل، اقلیم، موجودات زنده، پستی و بلندی، مواد مادری، زمان و موقعیت جغرافیایی برای پیش‌بینی مکانی متغیر در نظر گرفته می‌شود (McBratney et al, 2003). عوامل هفتگانه از منابع مختلف مانند مدل رقوم ارتفاع، داده‌های ماهواره‌ای و اطلاعات زمین‌شناسی محاسبه می‌شوند. پارامترهای اراضی که قابل استخراج از مدل رقوم ارتفاع هستند، نقش اساسی در بیان تغییرپذیری مکانی عملکرد محصولات دارند (Florinsky et al, 2004). ارتباط معنی‌دار پارامترهای اراضی از قبیل ارتفاع، شیب و انحنای با عملکرد گندم گزارش شده است (Kravchenko and Bullock, 2000). از داده‌های ماهواره‌ای اولین بار در پیش‌بینی عملکرد گندم در سال ۱۹۸۱ استفاده شده است (Tucker et al, 1981). Rembold et al (2013) نیز کارایی بالای تصاویر ماهواره را در تخمین پتانسیل عملکرد محصولات زراعی تأیید نمودند. ایشان ارتباط معنی‌داری را بین شاخص گیاهی نرمال شده و عملکرد محصولات گزارش کردند. شاخص گیاهی نرمال شده محبوبترین شاخص ماهواره‌ای برای ارزیابی کیفیت و کمیت محصولات معرفی شده است (Tucker et al, 1981; Macdonald and Hall, 1980). برای ارتباط دادن داده‌های کمکی و داده‌های واقعی زمینی (عملکرد گندم) از مدل‌های گوناگون از جمله برنامه‌ریزی ژنتیک می‌توان استفاده نمود. برنامه‌ریزی ژنتیک که برای اولین بار توسط Koza (1999) ارائه شد، جزو روش‌های الگوریتم‌گردشی محسوب می‌شود. مبنای تمامی این روش‌ها

گردیده است. بعضی زارعین مدیریتشان خوب بوده و مشکلی از نظر تأمین کود و آب و مدیریت خوب نداشته و عملکردشان بسیار خوب بوده در حالی که عملکردهای پایین مربوط به مدیریت ضعیف و محدودیت‌های خاکی بوده که موجب کاهش عملکرد شده است. به عنوان مثال میزان مصرف کود ازته دارای محدوده ۱۰۰ الی ۲۵۰ کیلوگرم در هکتار است که از طریق سرک در دو مرحله به خاک اضافه می‌شود. از آنجایی که اراضی تحت شبکه آبیاری و زهکشی است بنابراین محدودیتی از نظر تأمین آب وجود ندارد.

Typic Calcicusterts, Typic Haplusterts و پانزده فامیل قرار گرفتند (Soil Survey Staff, 2010). میزان عملکرد محصول گندم آبی در واحد سطح، در محل‌های حفر پروفیل اندازه‌گیری شد. در این منطقه ۱۱۰ نقطه بررسی گردید. نمونه‌برداری بصورت شبکه در ابعاد ۵۰۰ در ۵۰۰ متر انجام شد (شکل ۱). خصوصیات اندازه‌گیری شده خاک‌ها شامل pH، کربن آلی، آهک، ظرفیت تبادل کاتیونی، درصد رس، سیلت، ماسه، درصد سدیم قابل تبادل است. شکل (۲) نیز به طور خلاصه روند اجرای مطالعه حاضر را نشان می‌دهد. عملکردهای فعلی زارعین تحت مدیریت‌های مختلفی که زارع در شرایط فعلی داشته ثبت



شکل ۱- موقعیت مکانی محل نمونه‌برداری خاک و نمونه‌های اندازه‌گیری شده عملکرد گندم



شکل ۲- فرآیند انجام مطالعه و تهیه نقشه تغییرات مکانی عملکرد گندم

مشقتات اولیه و ثانویه استخراج گردید که شامل: ارتفاع، شیب، جهت شیب، انحنای شیب، انحنای سطحی شیب، شاخص خیسی، شاخص همواری دره با درجه تفکیک بالا، عمق دره، سطح حوزه اصلاح شده، سطح پایه شبکه آبراهه، شیب حوزه،

داده‌های کمکی و فرآیند انتخاب آن‌ها در مطالعه حاضر از انواع متغیرهای کمکی از جمله مدل رقومی ارتفاع و داده‌های ماهواره‌ای لندست استفاده گردید. از مدل رقومی تهیه شده از سازمان نقشه‌برداری با دقت مکانی ۱۰ متر،

زیرمجموعه‌ها را بر اساس تخمین دقت ناشی از انتخاب یک زیرمجموعه خاص، تحت شرایط الگوریتم طبقه‌بندی مورد استفاده انجام می‌دهد. الگوریتم‌های دسته دوم معمولاً نتایج بهتری به دست می‌دهند. مهم‌ترین بخش در روش انتخاب ویژگی حلقه بسته، الگوریتم جستجویی است که در آن به کار رفته است. در این مقاله از روش رپر با الگوریتم جستجوی تکاملی از نوع الگوریتم ژنتیک استفاده گردید. این روش به جعبه سیاه معروف است. دلیل استفاده از الگوریتم ژنتیک این است که این الگوریتم می‌تواند یک جستجوی تصادفی را انجام دهد و مستعد گیر افتادن در کمینه محلی نیست (Nosrati and Eftekhari, 2014). همچنین در تحقیق حاضر از کل اطلاعات کمکی برای مدل‌سازی استفاده گردید (مدل ۱) و نتایج آن با نتایج مدلی (مدل ۲) که تنها از لایه‌های رقومی انتخاب شده توسط روش رپر بهره گرفته بود، مقایسه گردید.

مدل استنباط مکانی و برنامه‌ریزی ژنتیک

برنامه‌ریزی ژنتیک جزو روش‌های الگوریتم گردشی محسوب می‌شود. مبنای این روش‌ها بر اساس نظریه تکاملی داروین استوار است. الگوریتم‌های یاد شده اقدام به تعریف یک تابع هدف در قالب معیارهای کیفی نموده و سپس تابع هدف را برای اندازه‌گیری و مقایسه روش‌های مختلف حل کرده و در یک فرآیند گام به گام تصحیح ساختار داده‌ها به کار گرفته و در نهایت، روش حل مناسب را ارائه می‌نماید. در برنامه‌ریزی ژنتیک در ابتدا بلوک‌های موجود که شامل متغیرهای ورودی و هدف و نیز تابع ارتباط دهنده آن‌ها می‌باشند، تعریف گردیده و سپس ساختار مناسب الگو و ضرایب آن تعیین می‌شود. فرآیند گام به گام برنامه‌ریزی ژنتیک به صورت مراحل زیر است: (۱) یک جمعیت اولیه از توابع مرکب نشان دهنده الگوهای پیش‌بینی، به صورت تصادفی در نظر گرفته می‌شود (ایجاد کروموزوم‌ها). (۲) معرفی جمعیت اولیه (کروموزوم‌ها) به رایانه و ارزیابی هر یک از افراد (ژن) جمعیت مذکور با استفاده از توابع برآزش (شناسایی مؤثرترین افراد در ماهیت پدیده). (۳) انتخاب ژن‌های مؤثر به منظور تکثیر، جهش، جفت‌گیری و تولید مثل افراد جدید با صفات اصلاح شده (فرزندان). (۴) اعمال فرآیند توسعه‌ای تکراری بر روی فرزندان در هر تولید گام چهارم به تعداد معین و یا تا حصول بهترین پاسخ تکرار خواهد شد (Koza, 1999). برای اجرای این مدل از جعبه ابزار برنامه‌ریزی ژنتیک (GPLAB) در نرم‌افزار متلب استفاده شد (Mathworks, 2010).

جهت شیب حوزه و ارتفاع حوزه است. همچنین از تصویر ماهواره‌ای پس از اصلاحات لازم (هندسی و رادیومتری) نسبت-های باندی مختلف از قبیل شاخص گیاهی نرمال شده، شاخص نسبت گیاهی، شاخص گیاهی تعدیل کننده اثر خاک، شاخص گچ، شاخص آهک، شاخص روشنایی، شاخص شوری و شاخص رس محاسبه گردید (جدول ۱). با استفاده از تجزیه مؤلفه‌ها تعداد باندهای تصویر (باند یک تا هفت) کاهش داده شد. تجزیه مؤلفه‌های اصلی^۱ یک روش- آماری چند متغیره است که متغیرهایی را که با هم همبستگی دارند بصورت مؤلفه‌های اصلی نشان می‌دهد. وقتی که متغیرها با هم همبستگی داشته باشند، تجزیه مؤلفه‌های اصلی برای کاهش داده‌های چند بعدی به تعداد کمتری از ترکیبات خطی متعامد، با خلاصه کردن منابع اصلی تغییرپذیری داده‌ها، مفید خواهد بود. به‌رحال اگر همبستگی بین متغیرها وجود نداشته باشد تعداد مؤلفه‌های اصلی تقریباً برابر متغیرها خواهد بود (Mouser et al, 2005). در انجام فرآیندهای تجزیه مؤلفه‌های اصلی، مؤلفه‌هایی که مقدار بردار ویژه^۲ آن‌ها بیشتر از یک باشد نگهداری می‌شوند؛ زیرا آن‌ها بیشترین تغییرات کل متغیرها را در خود حفظ می‌کنند (Khattree and Niak, 2000). و مؤلفه‌هایی که مقدار بردار ویژه آن‌ها کمتر از یک باشد حذف می‌شوند. بنابراین مؤلفه اصلی اول (PC1) بیشترین مقدار تغییرات را در کل متغیرهای اصلی نشان می‌دهد و به همین ترتیب مؤلفه‌های بعدی بیشترین تغییرات را نشان می‌دهند (Li et al, 2007). در مجموع ۲۴ لایه اطلاعات کمکی از منطقه مورد مطالعه تهیه گردید.

در اغلب موارد بسیاری از ویژگی‌ها (لایه‌های اطلاعاتی)، نامربوط و زائد هستند و کارایی الگوریتم یادگیری را کاهش داده و در بعضی موارد دقت و سرعت مدل‌سازی را پایین می‌آورند. بنابراین انتخاب لایه‌های اطلاعاتی مؤثر در فرآیند پیش‌بینی حائز اهمیت است. بسیاری از روش‌ها برای انتخاب ویژگی‌ها در دهه اخیر معرفی شده‌اند. الگوریتم‌های انتخاب ویژگی بسته به روند ارزیابی آن‌ها به دو دسته عمده تقسیم می‌شوند. اگر انتخاب ویژگی مستقل از هر گونه الگوریتم یادگیری انجام شود (یعنی به صورت پیش پردازنده کاملاً مجزا)، آن را روش فیلتر یا حلقه باز می‌گویند. در این مورد ویژگی‌های نامطلوب قبل از استنتاج دور ریخته می‌شوند. اما، اگر روند ارزیابی با یک الگوریتم طبقه‌بندی در ارتباط باشد، روش انتخاب ویژگی را روش رپر^۳ یا حلقه بسته می‌نامند. این روش، جستجو در فضای

1. Principal Component Analysis (PCA)
2. Eigenvalues
3. Wrapper

جدول ۱- داده‌های کمکی مورد استفاده جهت پیش‌بینی عملکرد گندم

تعریف	پارامترهای محاسبه شده
Elevation	ارتفاع بالای سطح دریای آزاد (متر)
Channel network based level	ارتفاع بالای شبکه زهکشی را نشان می‌دهد
Modified catchment area	تجمعات جریان آب و پارامترهای مربوطه
Valley depth	شاخص نسبت پهنای دره به عمق دره
Multi-Resolution valley bottom flatness index	میزان همواری و پائین بودن را محاسبه می‌کند
Wetness Index	معیاری از تجمع رطوبت
Slope	زاویه بین سطح تماس و سطح افقی در یک نقطه درسته زمین
Aspect	بیشترین سرعت تغییر در ارتفاع در هر سلول از مدل ارتفاعی
Catchment slope	میزان گرادیان بالای مسیر حرکت
Catchment aspect	میزان جهت شیب بالای مسیر حرکت
Catchment height	میزان ارتفاع بالای مسیر حرکت
Curvature	انحنای در صفحه عمودی خط جریان است.
Plan Curvature	بیان گر اندازه همگرایی و واگرایی جریان می‌باشد.
Principal component analysis	PC1, PC2, PC3
Normalized differential vegetation index	$(B4 - B3) / (B4 + B3)$
Soil-adjusted vegetation index	$[(B4 - B3) / (B4 + B3 + L)] \times [1 + L]$
Ratio vegetation Index	$B4 / B3$
Clay Index	$B5 / B7$
Gypsum Index	$(B5 - B4) / (B5 + B4)$
Salinity Index	$(B3 - B4) / (B2 + B4)$
Brightness Index	$((B3)^2 + (B4)^2) \cdot 0.5$
Carbonate Index	$B3 / B2$

بیشترین تغییرات و pH، سیلت، آهک به ترتیب با ۳/۸، ۸/۴ و ۱۵/۶ درصد کمترین تغییرپذیری را در منطقه داشتند. طبق طبقه‌بندی ارائه شده توسط Wilding (1985) چنانچه ضریب تغییرات کمتر از ۱۵ درصد باشد تغییرپذیری کم، بین ۱۵ تا ۳۵ درصد تغییرپذیری متوسط و چنانچه بیش از ۳۵ درصد گردد، تغییرپذیری زیاد است. حداقل عملکرد دانه گندم توسط زارع ۵۰۰ کیلوگرم در هکتار برای گندم و حداکثر آن ۵۷۰۰ کیلوگرم در هکتار است. همچنین حداقل هدایت الکتریکی ۱/۵ و حداکثر آن ۱۲۲/۳ دسی‌زیمنس بر متر است. البته مقادیر زیاد شوری مربوط به نواحی جنوبی منطقه مورد مطالعه است که فاقد هرگونه کشت و کاری است. جهت شناسایی عوامل خاکی تأثیرگذار بر عملکرد گندم در منطقه از روش رگرسیون گام به گام استفاده گردید. در این روش تمام خصوصیات خاک اندازه‌گیری شده شامل pH، کربن آلی، ظرفیت تبادل کاتیونی، شوری، قلبائیت، آهک، گچ، رس، سیلت و ماسه به عنوان متغیر مستقل و عملکرد کشاورز به عنوان متغیر وابسته وارد مدل شدند. نتایج حاصل نشان داد که تنها صفات شوری، گچ، pH و سیلت در سطح ۵ درصد ارتباط معنی‌داری با عملکرد داشته‌اند.

در نهایت جهت ارزیابی مدل‌ها، از روش حذف یک داده و پیش‌بینی آن استفاده گردید (Brus *et al*, 2011). در این روش، پایگاه داده (n) به n-1 موقعیت برای واسنجی و یک موقعیت برای اعتبارسنجی تقسیم می‌شود. در هر تکرار، مدل برای موقعیت حذف شده اجرا می‌گردد و متغیر حذف شده پیش‌بینی می‌گردد. سپس با مقدار واقعی مقایسه می‌گردد و خطای پیش‌بینی محاسبه می‌شود. این فرآیند برای همه موقعیت‌های نمونه-برداری انجام می‌شود. همچنین از معیارهای ارزیابی معمول از قبیل ریشه مربعات خطا، ضریب همبستگی تطابق لاین و ضریب تبیین در این مطالعه استفاده گردید (McBratney *et al*, 2003).

نتایج و بحث

خلاصه آماری داده‌ها

جدول (۲) خلاصه آمار خصوصیات فیزیکی و شیمیایی خاک-های منطقه مورد مطالعه و همچنین مهمترین متغیرهای کمکی مورد استفاده در این تحقیق را نشان می‌دهد. گچ، شوری، درصد سدیم قابل تبادل به ترتیب با ۲۶۸، ۲۲۳ و ۱۰۷ درصد

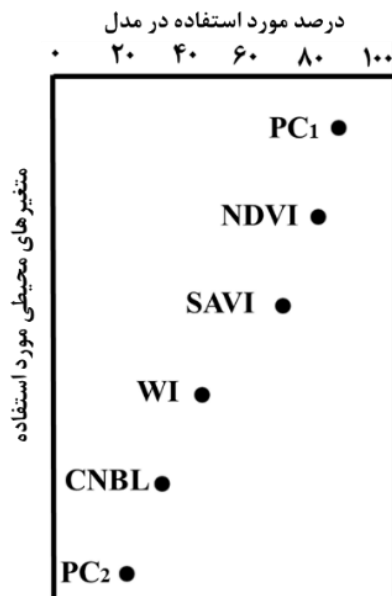
جدول ۲- آمار توصیفی خصوصیات فیزیکی و شیمیایی خاک در منطقه عقیلی گتوند و مهمترین متغیرهای کمکی

خصوصیات	واحد	حداقل	حداکثر	میانگین	انحراف معیار	ضریب تغییرات
pH	---	۶/۹	۸/۱	۷/۵۴	۰/۲۹	۳/۸
هدایت الکتریکی	dSm ⁻¹	۱/۵	۱۲۲/۳	۷/۳۳	۱۶/۳۷	۲۲۲/۳
کربن آلی	درصد	۰/۱۱	۱/۳	۰/۶۶	۰/۲۶	۳۹/۵
آهک	درصد	۳۵	۶۳	۴۷/۴۶	۷/۴۲	۱۵/۶
گچ	درصد	۰	۵/۶۲	۰/۳۵	۰/۹۳	۲۶۷/۹
ظرفیت تبادل کاتیونی	Cmolkg ⁻¹	۵	۲۱	۱۲/۰۹	۳/۷۹	۳۱/۳
رس	درصد	۱۱	۴۸	۳۰/۳۲	۸/۲۳	۱/۲۷
سیلت	درصد	۳۹	۵۹	۴۶/۶۹	۴/۰۷	۸/۴
شن	درصد	۷	۵۰	۲۱/۲۲	۸/۰۸	۳۸/۱
سدیم تبدالی	درصد	۰	۷۰	۱۱/۱۸	۱۱/۹۷	۱۰۷
عملکرد مشاهداتی	Kgha ⁻¹	۵۰۰	۵۷۰۰	۲۹۷۷/۶۳	۱۸۴۷/۲۰	۶۲/۰۳
مؤلفه اول تصویر	---	-۸/۲۴	۶/۳۸	-۶/۱۳	۲/۲۲	-۳/۶۳
مؤلفه دوم تصویر	---	-۲/۱۵	۲/۷۲	-۲/۸۲	۰/۸۲	-۲/۹۲
شاخص گیاهی نرمال شده	---	-۰/۳۲	۰/۶۷	-۰/۰۷	۰/۰۵	-۷۱/۳۴
شاخص گیاهی تعدیل کننده اثر خاک	---	-۰/۸۴	۱/۲۱	-۰/۱۱	۰/۰۸	-۷۱/۳۴
شاخص خیسی	---	۶/۹۱	۱۷/۹۳	۱۱/۷۰	۲/۵۶	۲۱/۹۴
سطح پایه شبکه آبراهه	متر	۵۵/۰۰	۶۵/۰۱	۶۰/۵۷	۱/۷۵	۲/۹۰

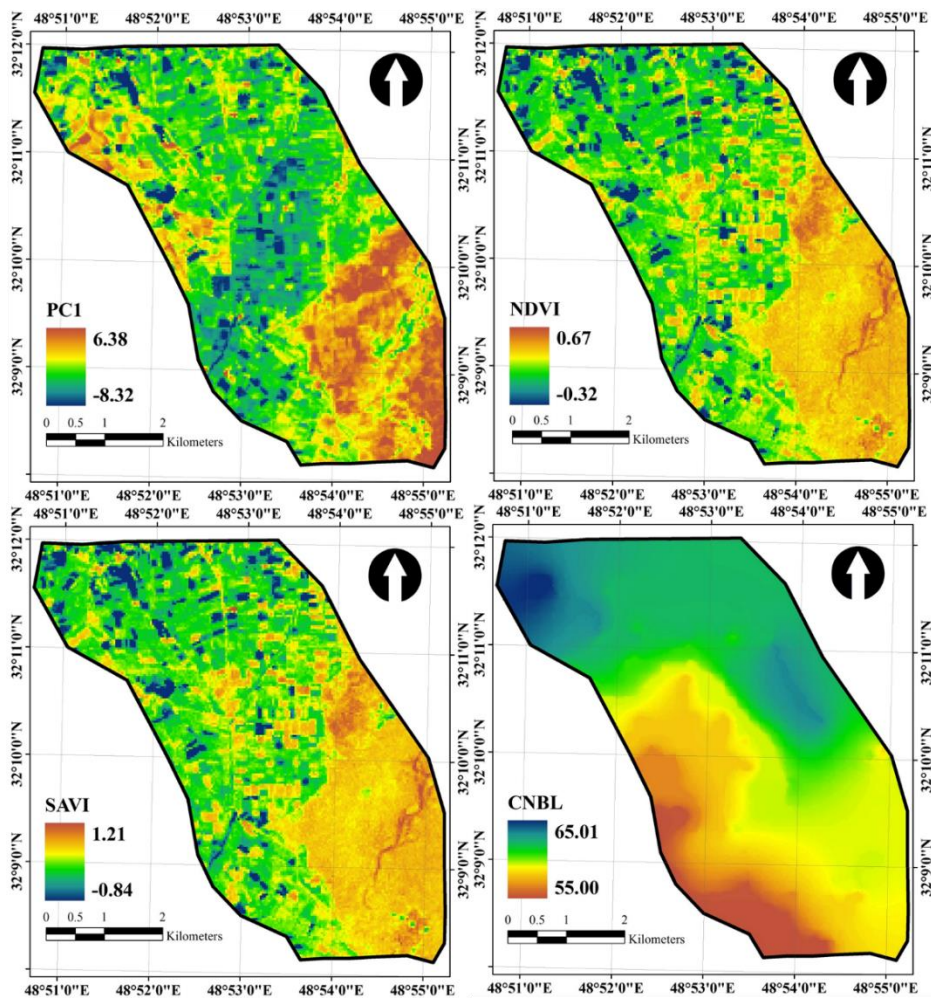
متغیرهای کمکی و انتخاب آن‌ها

نتایج الگوریتم رپر نشان داد که بعضی از متغیرهای کمکی دارای اهمیت بیشتری برای وارد شدن در فرآیند مدل‌سازی می‌باشند. مطابق با شکل (۳) ملاحظه می‌شود که تجزیه مؤلفه اول دارای بیشترین اهمیت است. پس از آن به ترتیب لایه‌های رقومی شاخص گیاهی نرمال شده، شاخص گیاهی تعدیل کننده اثر خاک، شاخص خیسی، سطح پایه شبکه آبراهه و تجزیه مؤلفه دوم قرار دارند. (Rembold et al 2013) نیز کارایی بالای تصاویر ماهواره را در تخمین پتانسیل عملکرد محصولات زراعی تأیید نمودند. ایشان ارتباط معنی‌داری را بین شاخص گیاهی نرمال شده و عملکرد محصولات گزارش کردند. Tucker et al (1983) نیز کارایی قابل قبول تصاویر ماهواره جهت پیش‌بینی تولید ماده خشک گیاهان را تأیید نمودند. لذا مدل برنامه‌ریزی ژنتیک با استفاده از دو دسته ورودی (همه ورودی‌ها و ورودی-های انتخاب شده توسط الگوریتم رپر) اجرا گردید. در شکل (۴) تغییرات مکانی بعضی از داده‌های کمکی مشاهده می‌شود. تجزیه مؤلفه اول دارای مقدار کمینه و بیشینه به ترتیب -۸/۲۴-

و ۶/۳۸ است. بیشترین مقادیر در قسمت‌های شمالی منطقه دیده می‌شود. شاخص گیاهی نرمال شده دارای محدوده ۰/۳۲- الی ۰/۶۷ است. بیشترین مقدار آن در نواحی شمالی یعنی در جایی که زمین‌های کشاورزی و پوشش گیاهی مناسب است دیده می‌شود. مقادیر بیش از ۰/۵ آن مربوط به پوشش گیاهی قوی است. به طور مشابه شاخص گیاهی تعدیل کننده اثر خاک نیز روند مشابه‌ای را نشان می‌دهد. به طوری که کمترین مقدار آن در نواحی فاقد پوشش گیاهی و جنوب منطقه مورد مطالعه بدست آمده است. سطح پایه شبکه آبراهه که از مدل رقومی ارتفاع محاسبه شده است دارای محدوده عددی بین ۵۵ الی ۶۵ است. بیشترین مقدار آن در قسمت‌های شمالی منطقه دیده می‌شود. (Shahbazi et al 2009) نیز ارتباط معنی‌داری را بین پارامترهای مستخرج شده از مدل رقومی ارتفاع و عملکرد گندم را تأیید نمودند. همچنین شاخص خیسی یکی از داده‌های کمکی مؤثر در پیش‌بینی عملکرد گندم شناخته شد که (Shabani et al 2012) نیز ارتباط معنی‌دار شاخص خیسی و عملکرد را گزارش کرده‌اند. اما در کل تغییرات ارتفاعی منطقه مورد مطالعه در این پژوهش بسیار ناچیز است.



شکل ۳- اهمیت متغیرهای کمکی منتخب توسط الگوریتم ربر (PC: مؤلفه‌های اول و دوم؛ NDVI: شاخص گیاهی نرمال شده؛ SAVI: شاخص گیاهی تعدیل کننده اثر خاک، WI: شاخص خیسی، CNBL: سطح پایه شبکه آبراهه)



شکل ۴- پراکنش مکانی بعضی از متغیرهای کمکی منتخب توسط الگوریتم ربر (PC: مؤلفه اول تصویر ماهواره‌ای؛ NDVI: شاخص گیاهی نرمال شده؛ SAVI: شاخص گیاهی تعدیل کننده اثر خاک؛ CNBL: سطح پایه شبکه آبراهه)

مدل استنباط مکانی و برنامه‌ریزی ژنتیک

جهت ارتباط داده‌های کمکی با عملکرد گندم دو رویکرد انتخاب گردید. زمانی که تمامی داده‌های کمکی به عنوان ورودی انتخاب شد (مدل ۱) و زمانی که فقط داده‌های کمکی منتخب توسط الگوریتم رپر وارد مدل گردید (مدل ۲). نتایج نشان داد که پارامترهای مدل برنامه‌ریزی ژنتیک (مدل ۱) به این ترتیب بهینه شدند: تولید بدون بهبود ۴۲۵، تولید از زمان شروع ۵۰۰، حداکثر مقدار اجرا ۳۷۰، حداکثر اندازه برنامه ۷۰۰، اندازه جمعیت ۴۶۰، نرخ جهش ۰.۹۵٪ و نرخ تلاقی ۰.۲۰٪، که رابطه بدست آمده در معادله (۱) ارائه شده است. در این مطالعه شش عملگر ریاضی شامل جمع، تفریق، ضرب، تقسیم، تون، رادیکال و توابع مثلثاتی جهت پیش‌بینی عملکرد گندم استفاده گردید. نتایج ارزیابی متقابل مدل (۱) نشان داد که آماره‌های میانگین ریشه مربعات خطا، ضریب همبستگی تطابق لاین و ضریب تبیین به ترتیب ۵۲۵/۱۱، ۰/۸۷ و ۰/۸۲ است. همچنین نتایج مدل برنامه‌ریزی ژنتیک (مدل ۲) که از داده‌های کمکی کمتری استفاده کرده بود نشان داد که پارامترهای بهینه شده به این صورت است: بدون بهبود ۳۲۲، تولید از زمان شروع ۲۵۰، حداکثر مقدار اجرا ۱۵۰، حداکثر اندازه برنامه ۴۲۵، اندازه جمعیت ۲۸۰، نرخ جهش ۰.۹۵٪ و نرخ تلاقی ۰.۲۰٪، که رابطه بدست آمده در معادله (۲) ارائه شده است. نتایج ارزیابی متقابل مدل (۲) نشان داد که آماره‌های میانگین ریشه مربعات خطا، ضریب همبستگی تطابق لاین و ضریب تبیین به ترتیب ۵۳۰/۸۲، ۰/۸۶ و ۰/۷۹ است (شکل ۵). نتایج ارزیابی مدل نسبت به نتایج Shahbazi et al (2009) بهتر است که دلیل آن می‌تواند مربوط به استفاده همزمان تصاویر ماهواره و مدل رقومی ارتفاع در پژوهش حاضر باشد. همچنین در تحقیق از مدل برنامه‌ریزی ژنتیک استفاده گردید که کارایی بالای آن در مطالعات خاکشناسی تأیید شده است (Johari et al, 2006; Makkeasorn et al, 2006; Parasuraman et al, 2007; Padarian et al, 2012). در کل با مقایسه نتایج دو مدل (۱ و ۲)

متوجه می‌شویم که مدل (۱) علاوه بر پیچیده‌تر شدن و زمان‌بر بودن (۲۳ ساعت و ۱۶ دقیقه)، آماره‌های میانگین ریشه مربعات خطا و ضریب تبیین را نتوانسته به صورت معنی‌داری (مقایسه دانکن) نسبت به زمانی که فقط از داده‌های انتخاب شده توسط مدل رپر استفاده گردید (مدل ۲)، افزایش دهد. این نتایج نشان می‌دهد که الگوریتم رپر به خوبی توانسته است متغیرهای مفید را انتخاب کند و داده‌های زائد را از فرآیند مدل‌سازی خارج کرده و رگرسیون ساده‌تر شده‌ای را ارائه نماید. بنابراین با استفاده از معادله (۲) میزان عملکرد برای کل منطقه مورد مطالعه تخمین زده شد که نتایج در شکل (۶) دیده می‌شود. لازم به ذکر است که در جنوب شرقی منطقه به علت وجود رودخانه شور و شوری خیلی زیاد خاک هیچگونه کشت گندم آبی انجام نشده و کشت آبی گندم در مناطق شمال غربی و شرقی که دارای شبکه آبیاری و زهکشی است انجام شده و عملکرد در این نواحی خوب است و بطور کلی عملکرد گندم آبی از سمت جنوب به شمال به علت کاهش شوری افزایش می‌یابد.

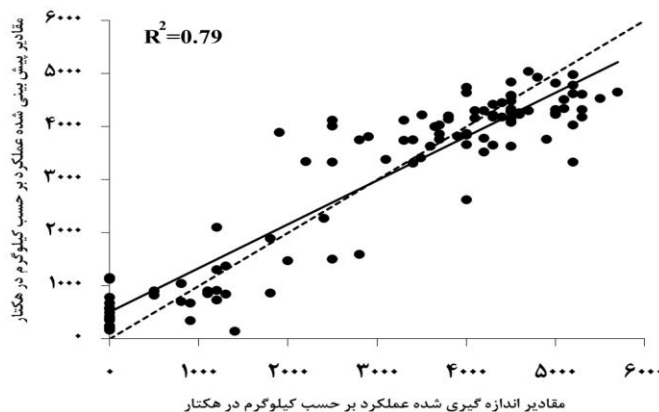
$$Y = \frac{-10.8}{DEM-54.8} + \frac{\cos(0.2CH)}{0.004(CS+Slope^2 \times 0.2(CA+0.003Aspect))} + 0.1 \cos(0.26 + PC1 + PC2 + PC3)$$

$$-48RVI - 4.24SI \cos BI^2 - 0.13 \sin CI -$$

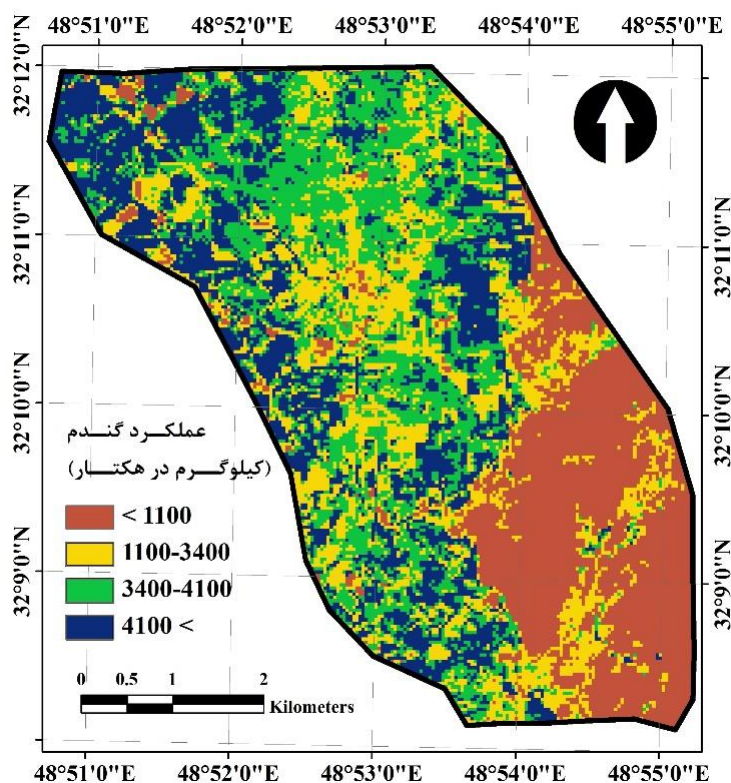
$$424 \sin(0.9 \cos CaI) \cos(0.06GI) - 0.01NDVI^2 \sin(SAVI) \sin(0.09Cu) - (0.41CNBL \cos(WI^2)) - 0.2 \frac{MrVBF}{VD}$$

$$Y = \frac{-16.1}{CNBL-60.6} + \frac{0.57WI-34.6}{0.31-PC1} + 981 \cos(0.37 + PC1 - 1.44PC2) - 0.001PC1 - 0.03SAVI - 0.1NDVI^2$$

که در آن Y: عملکرد گندم؛ PC: مؤلفه‌های اول، دوم و سوم تصویر ماهواره‌ای؛ CH: ارتفاع حوزه؛ CS: شیب حوزه؛ Slope: شیب؛ CA: جهت شیب حوزه؛ RVI: نسبت شاخص گیاهی؛ SI: شاخص شوری؛ BI: شاخص روشنایی؛ CI: شاخص رس؛ CaI: شاخص کربنات؛ GI: شاخص گچ؛ NDVI: شاخص گیاهی نرمال شده؛ Cu: انحنای شیب؛ PrCu: انحنای طولی شیب؛ MrVBF: شاخص همواری دره با درجه تفکیک بالا؛ SAVI: شاخص گیاهی تعدیل کننده اثر خاک، WI: شاخص خیسی، CNBL: سطح پایه شبکه آبراهه).



شکل ۵- مقادیر اندازه‌گیری شده و پیش‌بینی شده عملکرد گندم (کیلوگرم در هکتار) با استفاده از مدل برنامه‌ریزی ژنتیک (مدل ۲)



شکل ۶- پراکنش مکانی عملکرد گندم (کیلوگرم در هکتار) با استفاده از مدل برنامه‌ریزی ژنتیک (مدل ۲)

نتیجه مدل رگرسیونی بسیار ساده‌تری ارائه دهد. همچنین مدت زمان اجرای برنامه جهت رسیدن به جواب بهینه را نیز کاهش دهد. نتایج تحقیق همچنین کارایی بالای برنامه‌ریزی ژنتیک را در پیش‌بینی عملکرد گندم تأیید می‌کند. لذا پیشنهاد می‌شود در تحقیقات آینده از ترکیب الگوریتم رپر و برنامه‌ریزی ژنتیک جهت برآورد مکانی عملکرد سایر محصولات استفاده شود. همچنین پیشنهاد می‌گردد از داده‌های کمکی دیگر بروی تغییرات مکانی عملکرد محصولات تأثیرگذارند نیز استفاده شود.

نتیجه‌گیری

در پژوهش حاضر از مدل برنامه‌ریزی ژنتیک جهت بررسی تغییرات مکانی عملکرد گندم در منطقه گتوند استان خوزستان استفاده گردید. داده‌های کمکی مورد استفاده، از دو منبع رایگان و سهل‌الوصول (مدل رقومی ارتفاع و داده‌های سنجد لندست) بدست آمد. همچنین در این تحقیق سعی شد تا کارایی الگوریتم رپر (الگوریتم انتخاب ویژگی) در حذف داده‌های کمکی نامربوط بررسی شود. نتایج بررسی نشان داد الگوریتم رپر به خوبی توانست داده‌های کمکی مفید را شناسایی کرده و در

REFERENCES

- Becker-Reshef, E., Vermote, A., Lindeman, M. and Justice, C. (2010). A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. *Remote Sensing of Environment*, 114, 1312–1323.
- Brus, DJ., Kempen, B. and Heuvelink, GBM. (2011). Sampling for validation of digital soil maps. *European Journal of Soil Science*, 62, 394–407.
- Cahn, M. D., Hummel, J. W. and Brouer, B. H. (1994). Spatial analysis of soil fertility for site-specific crop management. *Soil Science Society American Journal*, 58, 1240-1248.
- Charles, J., Godfray, J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M. and Toulmin, C. (2010). Food security: the challenge of feeding 9 billion people. *Science*, 327, 812–818.
- Chen, Z. X., Ren, J. Q., Zhou, Q. B. and Tang, H. J. (2008). Regional yield estimation for winter wheat with MODIS-NDVI data in Shandong, China. *International Journal of Applied Earth Observation and Geoinformation*, 10, 403–413.
- Chipanshi, A. C., Ripley, E. A. and Lawford, R. G. (1999). Large-scale simulation of wheat yields in a semi-arid environment using a crop-growth model. *Agricultural Systems*, 59, 57–66.
- Doraiswamy, P. C., Moulin, S., Cook, P. W. and Stern, V. (2003). Crop yield assessment from remote sensing. *Photogrammetric Engineering and Remote Sensing*, 69, 665–674.
- Florinsky, IV., McMahon, S. and Burton, DL. (2004). Topographic control of soil microbial activity: a case study of denitrifiers. *Geoderma*, 119: 33-53.
- Johari, A., Habibagahi, G. and Ghahramani, A. (2006).

- Prediction of soil-water characteristic curve using genetic programming. *Journal of Geotechnical and Geoenvironmental Engineering*, 132, 661-665.
- Johari, A., Habibagahi, G. and Ghahramani, A. (2006). Prediction of soil-water characteristic curve using genetic programming. *Journal of Geotechnical and Geoenvironmental Engineering*, 132, 661-665.
- Khattree, R. and Naik, D. N. (2000). Multivariate Data Reduction and Discrimination with SAS Software. SAS Institute Inc., Cary, NC.
- Koza, J., Bennett, H., Andre, D. and Keane, M. (1999). Genetic programming 400 III: Darwinian invention and problem solving. Burlington, MA: Morgan Kaufmann.
- Kravchenko, AN. And Bullock, DG. (2000). Correlation of corn and soybean grain yield with topography and soil properties. *Agronomy Journal*, 92, 75-83.
- Li, Y., Shi, Z., Li, F. and Li, H. Y. (2007). Delineation of site-specific management zones using fuzzy clustering analysis in a coastal saline land. *Computers and Electronics in Agriculture*, 56, 174-186.
- Macdonald, RB. and Hall, FG. (1980). Global crop forecasting. *Science*, 208, 670-679.
- Makkeasorn, A., Chang, NB., Beaman, M., Wyatt, C. and Slater, C. (2006). Soil moisture estimation in a semiarid watershed using RADARSAT- 1 satellite imagery and genetic programming. *Water Resources Research*, 42, 1-15.
- Maselli, F. and Rembold, F. (2001). Analysis of GAC NDVI data for cropland identification and yield forecasting in Mediterranean African countries. *Photogrammetric Engineering and Remote Sensing*, 67, 593-602.
- Mathworks, 2010. Matlab Version 7.0. The Mathworks Inc., Natick, MA.
- McBratney, AB., Mendonca-Santos, ML. and Minasny, B. (2003). On digital soil mapping. *Geoderma*, 117,3-52.
- Mouser, P. J., Rizzo, D. M., Roling, W. F. M. and Van Breukelen, B. M. (2005). A multivariate statistical approach to spatial representation of groundwater contamination using hydrochemistry and microbial community profiles. *Environmental Science & Technology*, 39, 7551-7559.
- Nosrati, H. and Eftekhari, M. (2014). A new approach for variable selection using fuzzy logic. *Computational Intelligence in Electrical Engineering*, 4, 71-83.
- Padarian, J., Minasny, B. and McBratney, A. (2012). Using genetic programming to transform from Australian to USDA/FAO soil particle-size classification system. *Australian Journal of Soil Research*, 50, 443-446.
- Padarian, J., Minasny, B. and McBratney, A. (2012). Using genetic programming to transform from Australian to USDA/FAO soil particle-size classification system. *Australian Journal of Soil Research*, 50, 443-446.
- Padilla, F. L. M., Maas, S. J., Gonz, M.P., Lez-Dugo., F. Mansilla, N., Rajan, Gavil, P., and Donguez, J. (2012). Monitoring regional wheat yield in Southern Spain using the GRAMI model and satellite imagery. *Field Crops Research*, 130, 145-154.
- Parasuraman, K., Elshorbagy, A. and Si, BC. (2007). Estimating saturated hydraulic conductivity using genetic programming. *Soil Science Society of American Journal*, 71, 1676-1684.
- Rembold, F., Atzberger, C., Savin, I. and Rojas, O. (2013). Using Low Resolution Satellite Imagery for Yield Prediction and Yield Anomaly Detection. *Remote Sensing*, 5, 1704-1733.
- Shabani, A., Haghnia, Gh., Karimi, A. and Ahmadi, M.M. (2012). Influence of Topography and Soil Characteristics on the Rainfed Wheat Yield in Sisab Region, Northeastern Iran. *Journal of Water and Soil*, 26, 922-932.
- Shahbazi, F., Jafarzadeh, A.A., Sarmadian, F., Neyshaboury, M.R., Oustan, Sh., Anaya-Romero, M. and De la Rosa, D. (2009). Suitability of Wheat, Maize, Sugar Beet and Potato Using MicroLEIS DSS Software in Ahar Area, North-West of Iran. *American-Eurasian Journal of Agriculture and Environment Science*, 5, 45-52.
- Soil survey staff. (2010). Keys to soil taxonomy. Eleventh Edition. USDA. NRCS. 338pp.
- Taghizadeh-Mehrjardi, R. (2015). Digital mapping of cation exchange capacity using genetic programming and soil depth functions in Baneh region, Iran. *Archive of Agronomy and Soil Science*, xxx-xxx.
- Toscano, P., Ranieri, R., Matese, A., Vaccari, F. P., Gioli, B. A., Zaldea, M., Silvestri, C., Ronchi, P., La-Cava, J.R., Porter, A. and Miglietta, F. (2012). Durum wheat modeling: The Delphi system, 11 years of observations in Italy. *European Journal of Agronomy*, 43, 108-118
- Tucker, C.J., Vanpraet, C., Boerwinkel, E. and Gaston, A. (1983). Satellite remote sensing of total dry matter production in the Senegalese Sahel. *Remote Sensing of Environment*, 13, 461-474.
- Tucker, C.J., Holben, B.N., Elgin, J.H. and McMurtrey, J.E. (1981). Remote sensing of total dry-matter accumulation in winter wheat. *Remote Sensing of Environment*, 11, 171-189.
- Wall, L., Larocque, D. and Leger, P. M. (2007). The early explanatory power of NDVI in crop yield modeling. *International Journal of Remote Sensing*, 29, 2211-2225.
- Wilding, L. (1985). Spatial variability. Its documentation, accommodation, and implication to soil surveys. In: D. R. Nielson and J. Bouma (Eds). *Soil Variability*, Pudo, Wagenigen, the Netherlands.