

تخمین عملکرد گندم آبی با استفاده از روش الگوریتم ترکیبی شبکه‌های عصبی مصنوعی و الگوریتم ژنتیک

علی باریکلو^۱ - پریسا علمداری^{۲*} - کامران مروج^۳ - مسلم ثروتی^۴

تاریخ دریافت: ۱۳۹۵/۰۳/۲۶

تاریخ پذیرش: ۱۳۹۵/۱۲/۱۷

چکیده

از اهداف کشاورزی پایدار افزایش راندمان تولید محصولات کشاورزی با اعمال مدیریت‌های صحیح بوده و لازمه آن درک کامل‌تر روابط بین میزان تولید محصول با ویژگی‌های خاک و محیط می‌باشد. نخستین قدم یافتن روش‌های مناسبی است که توانایی تعیین روابط صحیح بین ویژگی‌های اراضی با مقدار عملکرد محصول باشد. هدف از این مطالعه بررسی کارایی مدل ترکیبی ژنتیک-عصبی در برآورد عملکرد گندم آبی در غرب شهرستان هریس می‌باشد. منطقه مطالعاتی در شمال شرق تبریز واقع شده و رژیم حرارتی و رطوبتی خاک به ترتیب مزیک و اریدیک هم مرز با زیریک می‌باشد. گندم، هندوانه و یونجه از مهم‌ترین محصولات زراعی منطقه است. بدین منظور تعداد ۸۰ خاک‌رخ در مزارع گندم انتخاب و از هر افق ژنتیکی نمونه خاک اخذ و به آزمایشگاه منتقل و تجزیه‌های فیزیکی و شیمیایی روی نمونه‌ها صورت گرفت. جهت مدل‌سازی لایه‌های ورودی شامل ویژگی‌های شیمیایی، فیزیکی، زمین‌نما و خروجی عملکرد مشاهده شده گندم آبی می‌باشد. نتایج آنالیز حساسیت نشان داد که نیتروژن کل، فسفر قابل جذب، شیب، درصد سنگریزه، واکنش خاک و ماده آلی به عنوان ویژگی‌های اراضی مهم در عملکرد گندم در اراضی مورد مطالعه هستند. نیتروژن کل خاک به‌عنوان موثرترین ویژگی در کیفیت و کمیت عملکرد گندم بر اساس ماتریس همبستگی پیرسون ایجاد شده بین ویژگی‌ها و عملکرد می‌باشد. کارایی مدل ژنتیک-عصبی با موفقیت برای تشریح رابطه بین عملکرد گندم و ویژگی‌های زودبافت صورت گرفت، به طوری که دارای ضریب تبیین بالا (۰/۸۷) و میانگین انحراف مربعات خطا کم (۴۷۳/۵) بود. نهایتاً می‌توان نتیجه گرفت، مدل هیبریدی می‌تواند به‌عنوان یک ابزار قدرتمند در تخمین عملکرد گندم باشد.

واژه‌های کلیدی: کشاورزی پایدار، مدل‌سازی، مدیریت، ویژگی‌های دیربافت و زودبافت

مقدمه

می‌باشد. از طرفی خاک به‌عنوان یک سامانه حامی حیات، نیازمند پی‌گیری مستمر در شناخت آن و جمع‌آوری اطلاعات و داده‌های به هنگام و دقیق است. بسط و توسعه پایگاه‌های داده‌ها و اطلاعات خاک، مدیریت بهینه سامانه اطلاعاتی، به‌کارگیری فن‌آوری‌های نوین جمع‌آوری، ذخیره و پردازش اطلاعات (سنجش از دور، سامانه اطلاعات جغرافیایی، سامانه‌های هوشمند تصمیم‌گیرنده، مدل‌سازهای ریاضی و آماری، شبیه‌سازی و ...)، از نیازمندی‌های اساسی می‌باشند (۲۱). بنابراین به‌نظر می‌رسد برآورد عملکرد به‌عنوان یک پارامتر دیربافت از روی پارامترهای زودبافت از اهمیت فراوانی برخوردار باشد. متداول‌ترین روش برای پیش‌بینی میزان عملکرد از روی ویژگی‌های خاک استفاده از ارزیابی کمی تناسب اراضی به روش فائو است (۳۶ و ۳۷). مطالعات زیادی در این ارتباط در دنیا و ایران صورت گرفته است (۱، ۴، ۱۴ و ۱۵). از سایر روش‌های مورد استفاده جهت برآورد عملکرد برای یک محصول خاص می‌توان به مدل آبرو دلاروزا و همکاران (۸)، ثروتی و همکاران (۳۰)، مدل واترکراپ آلن و

امروزه در عرصه کشاورزی جهان، واژه کشاورزی پایدار دارای جایگاه ارزنده و ویژه‌ای است. کشاورزی پایدار، فعالیتی علمی و مبتنی بر اصول اکولوژیک است که هدف اصلی آن ایجاد حالت تعادل و رسیدن به پایداری در تولید می‌باشد (۳۹). همچنین مهم‌ترین مسئله برای کشاورزان یک منطقه کسب درآمد و نیل به حداکثر محصول است (۵). بنابراین می‌توان گفت افزایش تولید در واحد سطح با حفظ حاصلخیزی اراضی برای آیندگان امری اجتناب‌ناپذیر است. این امر نیازمند درک کامل روابط بین مقدار عملکرد با ویژگی‌های اراضی

۱، ۲ و ۳ - به ترتیب دانش آموخته کارشناسی ارشد و استادیاران گروه خاکشناسی، دانشکده کشاورزی، دانشگاه زنجان

*-نویسنده مسئول: (Email: p_alamdari@znu.ac.ir)

۴- استادیار مرکز آموزش میان‌دو، دانشگاه ارومیه

رده‌بندی آمریکایی (۲۰۱۴) در رده‌های اتی‌سول و اریدی‌سول قرار گرفتند (۳۴). همچنین برای تعیین میزان عملکرد در هر خاکرخ از پلات‌های مربعی چوبی ۱×۱ متر استفاده شد.

مطالعات آزمایشگاهی

آزمایش‌های فیزیکی و شیمیایی از جمله بافت به روش هیدرومتر (۱۳)، ظرفیت تبادل کاتیونی (CEC) به روش سابق و همکاران (۲۸)، کربن آلی (OC) به روش نلسون و سامرز (۲۳)، واکنش خاک در گل اشباع (۲۰)، کربنات کلسیم معادل (CCE) به روش تیتراسیون برگشتی (۲۴)، ازت کل به روش کج‌لدال (۷) و فسفر قابل دسترس به روش آبی (۲۷) بر روی نمونه‌های خاک صورت پذیرفت.

مدل‌سازی عملکرد گندم با مدل ترکیبی ژنتیک-عصبی

شبکه‌های عصبی مصنوعی تقلید بسیار ساده‌ای از مغز انسان، بر اساس یک پیکربندی ریاضی می‌باشد. به گونه‌ای که از چند لایه و در هر لایه از چند گره (نرون) تشکیل شده است. معمولی‌ترین نوع شبکه عصبی مصنوعی، مجموعه‌ای از نرون‌های پایه هستند که تشکیل دهنده لایه ورودی، یک یا چند لایه پنهان و یک لایه خروجی می‌باشند. داده‌های ورودی در خلال شبکه و در مسیری رو به جلو به صورت لایه به لایه منتشر می‌شود. این نوع شبکه عصبی، پیش‌تغذیه کننده یا پرسپترون چند لایه (MLP) نامیده می‌شود (شکل ۲). شبکه‌های عصبی در زمینه‌های مختلف به کار گرفته شده است. با این وجود، دستیابی به نتایج مطلوب تنها زمانی ممکن بوده است که منبع غنی داده‌ها با مشاهدات فراوان در اختیار باشد. زمانی که با داده‌های اندک مواجه شود، کارایی بالایی نداشته و کاربرد آن محدود می‌شود (۲). از سوی دیگر شبکه‌های عصبی مصنوعی فاقد یک روش استقرایی جامع در انتخاب الگوریتم یادگیری شبکه و توقف الگوریتم در حداقل محلی^۱ می‌باشد.

الگوریتم ژنتیک نیز جستجو را با جمعیتی از راه‌حل‌های اولیه تصادفی آغاز می‌کند. چنانچه معیارهای توقف برآورد نشود سه عملگر تکثیر، جهش و تقاطع به کار گرفته می‌شود تا جمعیت به روز شود. از آنجایی که نمایش راه‌حل‌ها در این الگوریتم بسیار شبیه به کروموزوم است و همچنین به دلیل شباهت عملگرهای آن به عملگرهای ژنتیکی، روند فوق الگوریتم ژنتیک خوانده می‌شود (۱۰). در این روش در ابتدای فرآیند هیچ‌گونه رابطه تابعی در نظر گرفته نشده و این روش قادر به بهینه‌سازی ساختار مدل و مؤلفه‌های آن می‌باشد. در استفاده از الگوریتم ژنتیک به دلیل تحمیل فرم تبعی خاص به صورت خطی، درجه دوم، نمایی و یا فرم دیگر، امکان شناسایی روابط پیچیده

همکاران (۳) و مدل واگنینگن اعتدالی و همکاران (۱۲) اشاره نمود. از روش‌های دیگر می‌توان روش‌های فراکوشی را نام برد. تقی زاده مهرجردی (۳۸) از مدل برنامه‌ریزی ژنتیک جهت مدل‌سازی عملکرد گندم آبی استفاده کردند و استفاده از یک ساختار مشخص در مدل‌سازی را عاملی در کاهش دقت مدل عنوان نمودند، با این حال گزارش نمودند که می‌تواند با دقت قابل قبولی عملکرد را پیش‌بینی نماید. سادات و همکاران (۳۵) نیز از شبکه‌های عصبی مصنوعی جهت برآورد پتانسیل تولید تیپ بهره‌وری انتخابی استفاده نمودند و گزارش کردند که دقت بیشتری از روش‌های خطی داشته ولی با این حال توقف یادگیری شبکه‌های عصبی در نقطه بهینه محلی باعث کاهش دقت مدل می‌شود. بنابراین برآورد عملکرد به‌عنوان یک ویژگی دیرپافت از روی مدل‌های خطی براساس نتایج تحقیقات همراه با عدم قطعیت، محاسبات زیاد و اطلاعات دقیق خاک، زمین‌نما، اقلیم، گیاه و مدیریتی است (۲۲، ۳۰) که روش‌های فراکوشی از جمله برنامه‌ریزی ژنتیک، شبکه‌های عصبی مصنوعی، نظریه مجموعه‌های فازی و ترکیب آن‌ها می‌تواند کارشناسان کشاورزی را در راستای رسیدن به این هدف یاری رساند. از طرفی این مدل‌ها نیز دارای یک سری ایراداتی هستند که انتظار می‌رود روش‌های ترکیبی محققان را در این راستا راهنما باشد.

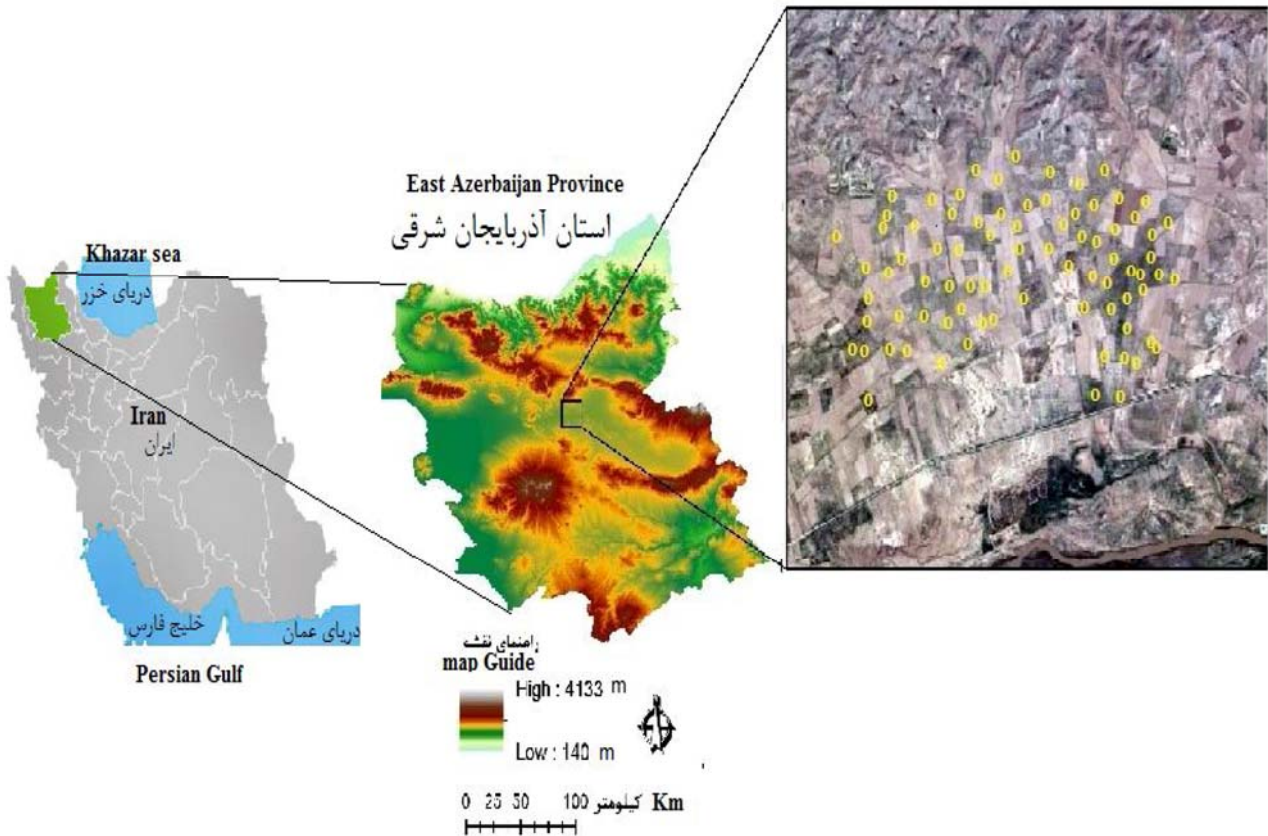
هدف از این تحقیق تخمین عملکرد گندم آبی با استفاده از مدل ترکیبی عصبی-ژنتیک و سنجش کارایی آن می‌باشد. کاربرد روش‌های فراکوشی ترکیبی جهت استفاده از مزایا و کاهش تاثیر معایب هر مدل و ارائه یک مدل مناسب در تخمین عملکرد یک محصول از نوآوری‌های این تحقیق می‌باشد.

مواد و روش‌ها

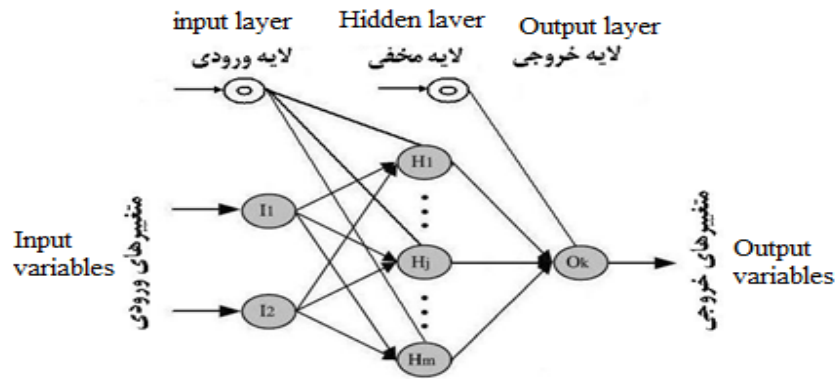
منطقه مورد مطالعه در استان آذربایجان شرقی و شمال شرق تبریز، در شهرستان هریس واقع شده است. این منطقه در مختصات جغرافیایی ما بین ۳۸ درجه و ۷ دقیقه تا ۳۸ درجه و ۱۱ دقیقه عرض شمالی و ۴۶ درجه و ۳۷ دقیقه تا ۴۶ درجه و ۴۴ دقیقه طول شرقی (شکل ۱) واقع شده است. از نظر زمین‌شناسی منطقه از مجموعه تشکیلات مارنی - آهکی دوره‌های پلیوسن و میوسن و تشکیلات آهکی همراه با رسوبات دوران چهارم بر روی خاکستر آتشفشانی تشکیل شده است (۳۲).

خاک‌های این منطقه دارای رژیم رطوبتی Aridic border to Xeric و رژیم حرارتی Mesic بر اساس نرم‌افزار نیوهال می‌باشند (۲۵). جهت نیل به اهداف ذکر شده، تعداد ۸۰ خاکرخ در مزارع گندم آبی انتخاب و بر اساس روش نقشه برداری آمریکا تشریح و از هر افق نمونه‌برداری خاک انجام و به آزمایشگاه منتقل گردید (۲۹). پس از تجزیه و تحلیل نتایج، خاک‌ها بر اساس سیستم جامع

غیرخطی وجود ندارد (۶).



شکل ۱- موقعیت اراضی مورد مطالعه با نقاط نمونه‌برداری
Figure 1- Location of study area with sampling points



شکل ۲- نمایی از شبکه پرسپترون چند لایه (MLP)، (۱۹)
Figure 2- View of Multilayer Perceptron network (MLP), (19)

ژنتیک به آموزش شبکه عصبی پرداخته و به تخمین تولید گندم با استفاده از داده‌های ورودی پرداخته‌شد. نهایتاً کارایی آن‌ها از طریق آماره‌های ضریب تبیین (R^2) (رابطه ۱)، جذر میانگین مربعات خطا (RMSE) (رابطه ۲)، نش-ساتکلیف (E) (رابطه ۳) مورد ارزیابی قرار گرفت.

برای مدل‌سازی با روش ترکیبی ابتدا از ویژگی‌های خاک و زمین‌نمای خاک‌رخ‌های انتخابی، میانگین وزنی از عمق ۱۰۰ سانتی‌متری باتوجه به یکساله بودن محصول انتخابی به‌عمل آمد (۳۷). سپس ویژگی‌های اراضی به‌عنوان ورودی و عملکرد به‌عنوان خروجی مدل در نظر گرفته‌شد. برای این منظور به کمک الگوریتم

پسخور^۴ تقسیم می‌شوند. تفاوت آنها در این است که در شبکه‌های پسخور، حداقل یک سیگنال برگشتی از یک نرون به همان نرون یا نرون‌های همان لایه و یا لایه قبل وجود دارد. در بیشتر موارد، شبکه‌های عصبی پسخور می‌توانند بسیار مفید واقع شوند. اما در ۸۰ درصد کاربردها از شبکه‌های عصبی پیشخور (شکل ۳) استفاده می‌شود (۱۹). روش کار در شبکه‌های عصبی پرسپترون که به صورت پرسپترون‌های تک لایه و چند لایه موجود هستند و جزو شبکه‌های عصبی پیشخور طبقه‌بندی می‌شوند، به این صورت است که پرسپترون تک لایه تنها می‌تواند مسائل مجزای خطی را دسته‌بندی کند و برای مسائل پیچیده‌تر لازم است از تعداد بیشتری لایه استفاده کنیم. شبکه‌های پیشخور چند لایه از یک و یا تعداد بیشتری لایه‌های میانی تشکیل شده‌اند. این شبکه، یک شبکه کاملاً بهم مرتبط می‌باشد. چرا که هر نرون در یک لایه به تمامی نرون‌های لایه بعدی مرتبط می‌باشد. اگر بعضی از این ارتباطات وجود نداشته باشد شبکه، یک شبکه مرتبط ناقص است (۱۱).

الگوریتم یادگیری پس انتشار خطا^۵ (BP) از نوع یادگیری با نظارت است. در این روش، هنگامی که ورودی به شبکه اعمال می‌شود، جواب شبکه با جواب هدفی که از طرف ما برای شبکه تعیین شده، مقایسه می‌شود و سپس خطای یادگیری محاسبه شده و از آن برای تنظیم پارامترهای شبکه استفاده می‌گردد. به گونه‌ای که اگر مجدداً به شبکه همان ورودی اعمال شود، خروجی شبکه به جواب هدف نزدیکتر گردد. الگوریتم BP یک قانون یادگیری برای پرسپترون‌های چند لایه است. این الگوریتم بر پایه قانون دلتا و با استفاده از مجموع اندازه مجذور خطا برای نرون‌های خارجی طراحی گردیده است (۹). پارامترهای الگوریتم ژنتیک و نیز تعداد جمعیت و حداکثر تکرار فرآیند یادگیری به صورت جدول ۱ تعیین شد. شکل ۴ نیز مراحل اجرای الگوریتم ترکیبی را نشان می‌دهد. در شکل ۴ نمادهای RGA-NN، MLP-MM و MLP-LM به ترتیب نشان‌دهنده شبکه‌های عصبی مبتنی بر الگوریتم ژنتیک، شبکه‌های عصبی چند لایه پیشخور با الگوریتم یادگیری ممنوم و شبکه‌های عصبی چند لایه پیشخور با الگوریتم یادگیری لونیبرگ مارکواریت^۶ می‌باشد. در مدل ترکیبی عصبی - ژنتیک پس از آموزش، خروجی هر تکنیک به مقیاس واقعی برگردانده شده و کارایی مدل از طریق آماره‌های مختلف سنجیده می‌شود. هدف از به کارگیری الگوریتم ژنتیک بهینه‌سازی پارامتر وزن شبکه عصبی مصنوعی بود. بنابراین تابع هدف الگوریتم ژنتیک تابعی از نتایج آماری شبکه عصبی مصنوعی است (۱۶). برای آموزش ابتدا تعداد P (جمعیت هر نسل)

$$R^2 = \frac{[\sum_{i=1}^n (q_i - \bar{q}_i)(\hat{q}_i - \bar{\hat{q}}_i)]^2}{\sum_{i=1}^n (q_i - \bar{q}_i) \sum_{i=1}^n (\hat{q}_i - \bar{\hat{q}}_i)^2} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (q_i - \hat{q}_i)^2} \quad (2)$$

$$E = \frac{\sum_{i=1}^n (O_i - \bar{I}_i)^2}{\sum_{i=1}^n (O_i - \bar{O}_i)^2} \quad (3)$$

در این روابط، O_i مقدار عملکرد مشاهده شده (میانگین ۲ ساله)، I_i مقدار عملکرد برآورد شده توسط مدل شبکه‌های عصبی مصنوعی، \bar{O}_i میانگین مقادیر عملکرد مشاهداتی، \bar{I}_i میانگین مقادیر عملکرد محاسباتی و n تعداد داده‌ها می‌باشد.

در این تحقیق از شبکه‌های عصبی پرسپترون چند لایه استفاده شده است و به منظور انجام پیش‌بینی و گرفتن خروجی از شبکه، برای آموزش و هم‌گرایی سریع‌تر و دقت بیشتر شبکه ابتدا ورودی‌های آن با استفاده از تکنیک حداقل - حداکثر (رابطه ۴) استاندارد شده و به داده‌های نرمال (بی بعد) در بازه -۱ تا ۱ تبدیل شدند.

$$X_n = 2 \frac{Xr - X \min}{X \max - X \min} - 1 \quad (4)$$

در این رابطه X_{\min} ، X_{\max} ، X_n ، X_r به ترتیب نشان‌دهنده مقادیر واقعی، نرمال شده، حداکثر و حداقل داده‌های تحت بررسی است. پس از مرحله نرمال‌سازی، تصادفی نمودن داده‌ها انجام شد. نتیجه این مرحله، داشتن مجموعه‌ای از ورودی و خروجی‌ها می‌باشد که در آن دسته‌های ورودی - خروجی دارای نظام خاصی نیستند. پس از پایان تصادفی نمودن داده‌ها، میزان اطلاعاتی که باید در فرآیند آموزش شبکه استفاده شود، مشخص می‌شود. بر این اساس بخشی از داده‌ها برای آموزش (۸۰ درصد) و بخشی دیگر برای آزمون شبکه (۲۰ درصد) در نظر گرفته می‌شود. تعداد نرون‌های لایه پنهان از طریق سعی و خطا بین ۱ تا ۱۰ نرون برآورد گردید. در لایه پنهان از توابع فعال‌سازی تانژانت هذلولی^۱ (رابطه ۵) و زیگموئید^۲ (رابطه ۶) استفاده شده است.

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (5)$$

$$f(x) = \frac{1}{1 + e^{-x}} \quad (6)$$

در این روابط X مقدار داده است.

به طور کلی شبکه‌های عصبی به دو نوع شبکه‌های پیشخور^۳ و

4- Feed forward

5- Back Propagation

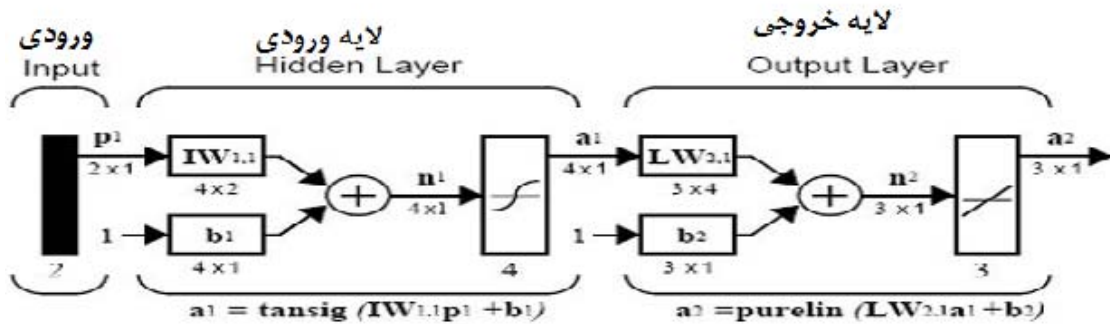
6- Levenberg-marquardt algorithm

1- Hyperbolic Tangent

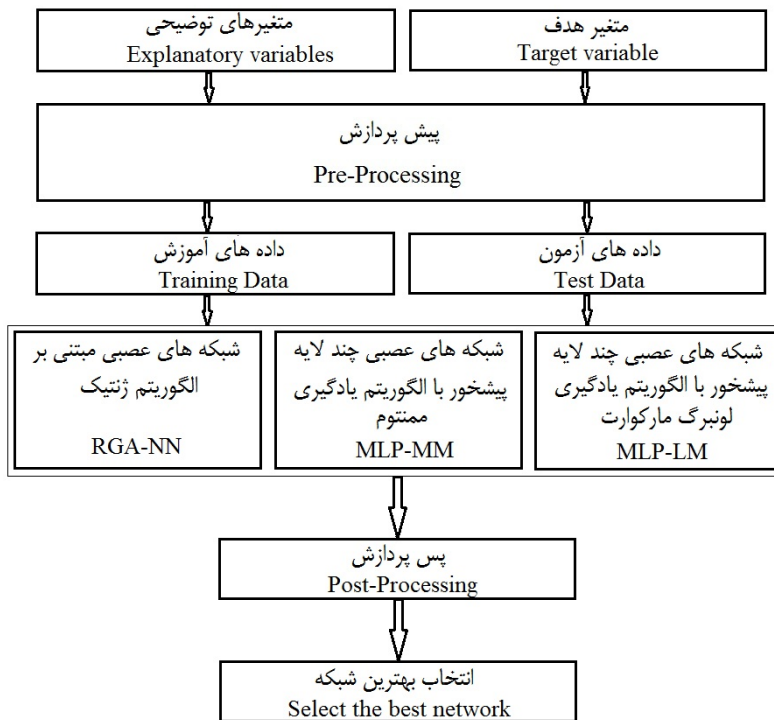
2- Sigmoid

3- Feedback

شبکه عصبی به طور تصادفی مقدار دهی اولیه شد و میزان خطا با استفاده از داده‌های آموزش محاسبه گردید.



شکل ۳- شبکه پیشخور دو لایه (۹)
Figure 3- Double layer feed forward networks (9)



شکل ۴- فلوچارت الگوریتم ترکیبی تخمین زنده عملکرد گندم آبی (۳۳)
Figure 4- Hybrid algorithm flowchart of wheat yield estimation (33)

جدید به دست آمده تکرار گردید. در مرحله آخر خروجی به دست آمده از اجرای شبکه با مقادیر واقعی مقایسه و در صورت حداقل شدن تفاوت بین دو مقدار بالا، اجرای مدل خاتمه یافت (۱۶).

در مرحله بعد به روزرسانی پارامترهای شبکه با توجه به مقادیر تولید مشاهده شده و نرخ‌های جهش^۱ و آمیزش^۲ انجام شد. الگوریتم بالا تا نبود بهبودی در پارامترهای شبکه با در نظر گرفتن جمعیت

نتایج و بحث

جدول ۱ تنظیم پارامترهای الگوریتم ژنتیک جهت مدل سازی و

- 1- Mutation
- 2- Cross over

علاوه بر معیارهای تعریف شده براساس جداول ساینس براساس نظرات کارشناسان محلی به عنوان داده‌های اولیه برای مدل‌سازی انتخاب شدند.

جدول ۲ پارامترهای آماری ویژگی‌های اراضی منتخب در مدل‌سازی عملکرد گندم آبی را در منطقه مطالعاتی نشان می‌دهد. ویژگی‌های انتخابی دارای دامنه وسیعی بوده و بر اساس جدول ساینس و همکاران (۳۶) در رشد و عملکرد گیاه گندم موثر هستند (۳۶). نیتروژن و فسفر

جدول ۱- تنظیم پارامترهای الگوریتم ژنتیک

Table 1- set the parameters of the genetic algorithm

تعداد جمعیت Population number	نرخ جهش Mutation Rate	نرخ تقاطع Intersection rate	حداکثر تکرار فرآیند یادگیری Maximum repeat of learning process
150	0.35	0.5	1000

جدول ۲- پارامترهای آماری ویژگی‌های اراضی انتخابی در مدل‌سازی عملکرد گندم

Table 2- Statistical parameters of selected land characteristics in wheat yield modeling

ویژگی‌های اراضی Land characteristics	حداقل Minimum	حداکثر Maximum	میانگین Average	ضریب تغییرات Coefficient of variation	واریانس Variance
شیب (درصد) Slope (%)	0	9	4.43	85.6	14.3
کربن آلی (درصد) Organic carbon (%)	0	1.1	0.8	48.3	0.154
pH	7.1	8.6	7.6	3.25	0.062
آهک (درصد) CaCO ₃	4.4	23.3	15.2	27.5	17.5
گچ (درصد) CaSo ₄	0.18	18.2	7.8	101.5	63.7
شن (درصد) Sand (%)	11.05	71.8	39.3	23.5	164.1
سیلت (درصد) Silt (%)	10.06	45.2	28.7	26.2	57.04
رس (درصد) Clay (%)	14.7	44.6	31.2	24.6	59.3
سنگریزه (درصد حجمی) Gravel(Volume percent)	0.43	37.1	8.39	82.4	48.09
CEC(Cmc.Kg ⁻¹)	0.73	34.1	17.7	26.3	19.42
نیتروژن کل (درصد) N (%)	4	16	9.59	34.2	10.7
فسفر(میلی‌گرم بر کیلوگرم) P(mg/k)	6.4	26.1	15.3	23.6	30.8
عملکرد واقعی (Kg ha ⁻¹) Real Yield	800.	6100	2916.25	561.36	1637074.36

در دانه‌ها تحت تأثیر میزان نیتروژن کل خاک بوده و یک پارامتر تأثیرگذار است. روش فائو نیز برای تعیین تناسب یک منطقه در مناطق خشک از کربن آلی استفاده نمی‌کند (۳۶) که نتایج تحقیق فوق نیز دلیل عدم استفاده از این پارامتر را در ارزیابی و تناسب اراضی یک منطقه خشک و نیمه‌خشک را برای محصول گندم تایید می‌نماید (۲۶).

جدول ۴ نتایج مقایسه سه مدل ترکیبی را با استفاده از پارامترهای

با توجه به ماتریس همبستگی پیروسون (جدول ۳) درصد شیب، درصد سنگریزه، نیتروژن، فسفر قابل دسترس، ماده آلی و pH به عنوان ورودی‌های مدل انتخاب شدند. نتایج موید این مطلب است که از بین ورودی‌های انتخابی، نیتروژن کل بیشترین و ماده آلی کمترین همبستگی را با عملکرد واقعی در منطقه نشان دادند. نوروژی و همکاران (۲۶) در پیش‌بینی عملکرد گندم در مناطق نیمه‌خشک غرب ایران بیان داشتند که میزان پروتئین موجود

آماري مختلف را نشان می‌دهد. نتایج نشان‌دهنده دقت بالای مدل الگوریتم لونبرگ مارکوارت نسبت به دو مدل دیگر است. این مدل دارای ضریب تبیین بالا، جذر میانگین مربعات خطای کمتر و ضریب نش - ساتکلیف نزدیک به ۱ دارد.

جدول ۳- ماتریس همبستگی بین پارامترهای انتخابی برای مدل‌سازی عملکرد گندم آبی
Table 3- correlation matrix of selected parameters for wheat yield modeling

عملکرد Yield	CEC	فسفر P	نیترژن N	سنگریزه Gravel	رس Clay	سیلت Silt	شن Sand	گچ CaSO ₄	آهک CaCO ₃	pH	کربن آلی Organic carbon	شیب Slop
شیب Slop												1
کربن آلی Organic carbon											0.6*	1
pH										1	0.83	0.003
آهک CaCO ₃									1	0.35	-0.90	0.16
گچ CaSO ₄								1	-0.16	-0.68	-0.18	0.22*
شن Sand							1	0.12	0.17	0.19*	0.19*	0.13
سیلت Silt						1	0.73*	0.03	0.21*	0.01	-0.13	0.22
رس Clay					1	0.35*	0.72*	-0.14	0.19*	0.22*	0.27*	-0.17
سنگریزه Gravel				1	0.44*	0.22*	0.44*	-0.67	0.05	0.19*	0.43*	0.32*
نیترژن N			1	0.30*	0.09	0.11	-0.14	-0.18	0.11	-0.14	0.17	0.23*
فسفر P		1	0.94*	0.29*	0.07	0.15	-0.14	0.26*	0.14	0.18	0.14	0.24*
CEC	1	0.21*	0.21*	0.44*	0.54*	0.12	0.45*	0.20*	0.1	0.13	0.42*	0.22*
عملکرد Yield	0.17	0.89*	0.91*	0.32*	0.07	0.11	-0.12	0.13	0.04	-0.12	0.19*	0.25*

* همبستگی در سطح احتمال ۰/۰۵ درصد

جدول ۴- پارامترهای آماری سه مدل تخمین عملکرد گندم آبی
Table 4- Statistical parameters of three wheat yield prediction models

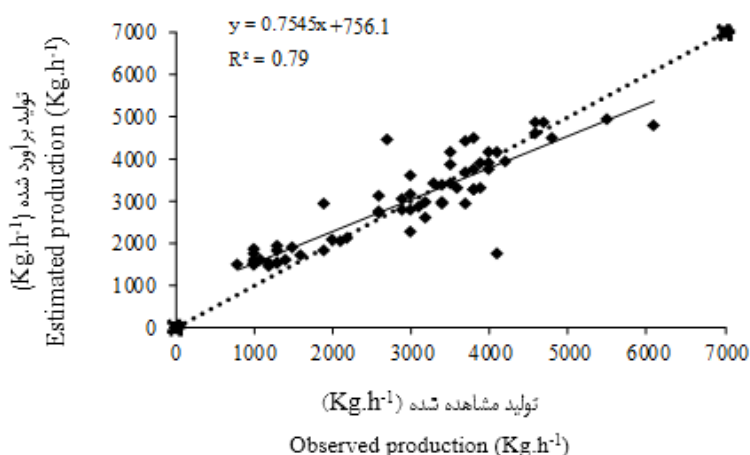
مدل Model	R ²	RMSE	NSE (E)
شبکه‌های مبتنی بر الگوریتم ژنتیک RGA-NN	0.87	473.5	-1.94
الگوریتم ممنوم MLP-MM	0.84	604.2	-2.84
الگوریتم لونبرگ مارکوارت MLP-LM	0.92	357.9	-1.23

۸۷ درصد تغییرات عملکرد بستگی به متغیرهای انتخابی و بقیه مربوط به عوامل کنترل نشده دارد. در ضمن میانگین انحراف مربعات خطا به ترتیب ۵۸۷ و ۴۷۳/۵ برای مرحله آموزش و آزمون برآورد گردید که با توجه به مقیاس عملکرد و تعداد داده مناسب است که انحراف کم

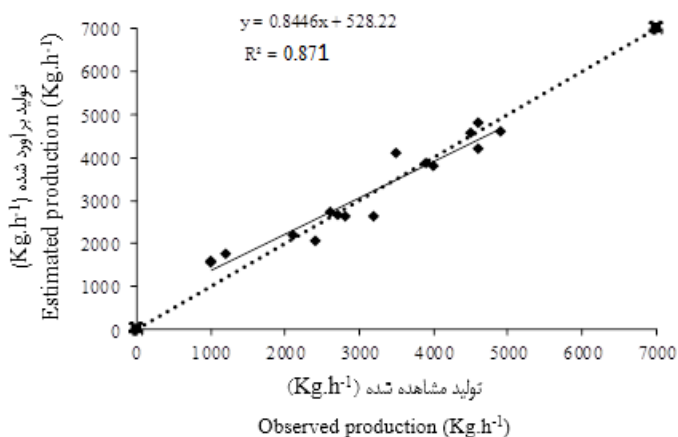
شکل ۵ همبستگی بین عملکرد واقعی و مشاهده شده در مرحله آموزش و شکل ۶ در مرحله آزمون را برای مدل انتخابی نشان می‌دهد. نتایج موید این مطلب است که روش ترکیبی دارای دقت بالایی بوده به طوری که در مرحله آموزش ۷۹ درصد و در مرحله آزمون

تقریباً منطبق بر نقاط واقعی هستند. این امر توانایی مناسب الگوریتم پیشنهادی را در تخمین عملکرد و در شرایط تقریبی محدودیت داده را نشان می‌دهد. لیو و همکاران (۱۸) و سادات (۳۵) در تحقیقات مشابهی تغییرپذیری مکانی ویژگی‌های اراضی را عامل ایجاد خطا در تخمین عنوان نمودند و پیشنهاد کردند واحدهای اراضی بایستی تا حد امکان کاهش یابد (۱۸ و ۳۵). در تحقیق حاضر در هریک از مزارع گندم خاکرخی حفر گردید و عملکرد نیز در ۵ نقطه از اطراف خاکرخی به روش پلات‌های ۱×۱ محاسبه و بعد از میانگین‌گیری به هکتار تعمیم داده شد. لازم به ذکر است که این کار در ۲ سال متوالی صورت گرفته است.

نمودار همبستگی تولید پیش‌بینی شده از تولید واقعی از خط نیمساز زاویه نیز در تایید این مطلب است. همچنین روند یادگیری شبکه براساس الگوریتم ژنتیک در انتهای فرآیند یادگیری خطای کمتری داشته است و الگوریتم ژنتیک در یافتن نقاط حداقل خطای مربوط به داده‌های آزمون (نقطه توقف و اوزان ایده‌آل شبکه)، این خطا را بسیار ناچیز ارائه می‌کند. نهایتاً نتایج موید این مطلب است که استفاده از تکنیک‌های غیرخطی در مدل‌سازی و تخمین عملکرد به‌عنوان یک پارامتر دیر یافت و غیرخطی با موفقیت همراه بوده است. همچنین مقادیر واقعی و پیش‌بینی‌شده عملکرد گندم براساس شبکه‌های عصبی مبتنی بر الگوریتم ژنتیک نشان می‌دهد که نقاط پیش‌بینی



شکل ۵- رابطه بین عملکرد مشاهده‌شده و پیش‌بینی شده در بخش آموزش
Figure 5- Relationship between observed and predicted yield in the train



شکل ۶- رابطه بین عملکرد مشاهده‌شده و پیش‌بینی شده در بخش آزمون
Figure 6- Relationship between observed and predicted yield in the test.

تحلیل فرم تبعی خاص از نقاط قوت الگوریتم ژنتیک و شبکه‌های عصبی می‌باشند. در این تحقیق یافته‌ها ضمن تایید نتایج مطلوب این تکنیک‌ها در صورت استفاده از تابع فعال‌سازی سیگموئیدی در لایه

نتیجه‌گیری کلی

الگوریتم ژنتیک و شبکه‌های عصبی از مهم‌ترین تکنیک‌ها در مدل‌سازی فرآیندهای غیرخطی بوده و جستجوی نقطه بهینه و عدم

پیشنهاد می‌شود. با توجه به دقت و صحت تخمین فوق استفاده از این تکنیک در تخمین سایر ویژگی‌های دیریافت خاک و گیاه که در تولید پایدار اراضی به‌عنوان شاخص می‌باشند، توصیه می‌گردد. نهایتاً می‌توان چنین عنوان کرد که معیارهای خاکی روابط بسیار قوی با میزان عملکرد محصول داشته، بنابراین مدیریت اراضی برای بهبود وضعیت ویژگی‌های خاک برای حفظ عملکرد بالا، جلوگیری از تخریب اراضی و حفظ آن برای آیندگان در راستای ایجاد یک کشت پایدار اجتناب‌ناپذیر است. همچنین در صورت وجود تغییرات ویژگی‌های خاک در طول زمان از جمله شوری می‌توان به نتایج دقیق‌تر و کاربردی‌تری در راستای توسعه پایدار کشاورزی با در نظرگیری تغییرات اقلیمی دست یافت.

پنهان و تابع فعال‌سازی خطی در لایه خروجی در کلیه شبکه‌های عصبی، نشان داد که تکنیک ترکیبی پیشنهادی نتایج بسیار بهتری داشته است. همچنین این نتایج قدرت پیش‌بینی بهتر شبکه‌های عصبی مبتنی بر الگوریتم پس‌انتشار خطا با تابع آموزشی لونیبرگ - ماکوارت در مقایسه با انواع دیگر شبکه‌های عصبی را تایید کرده و نشان می‌دهد که شبکه‌های عصبی مبتنی بر الگوریتم ژنتیک با کد گذاری حقیقی، با حداقل خطای تخمینی در داده‌های آموزش و آزمون به ترتیب برابر با ۰/۷۹ و ۰/۹۲ روش مناسب جهت تخمین عملکرد گندم در منطقه مطالعاتی می‌باشد. همچنین با توجه به تایید قدرت توضیح‌دهندگی متغیرهای خاک و زمین‌نمای انتخابی در این مطالعه و توصیف بهتر شرایط موثر بر عملکرد گندم در منطقه، استفاده از این متغیرها در مطالعات آتی جهت تخمین مطلوب‌تر تقاضای انرژی

منابع

1. Ashraf S.H., Ashraf V., and Abbaspour H. 2011. Assessment of land production potential for barley using geographic information system (GIS) method. *Indian Journal of Science and Technology*, 4(12):1775-1777.
2. Andonie R. 2010. Extreme Data Mining: Inference from small data sets. *International Journal of Computers Communications and Control*, 5(3): 280-291.
3. Allen R.G., Pereira, L.S Raes D., and Smith M. 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. FAO, Rome, 300(9):D05109.
4. Ayoubi SH.A Givi J., Jalalian A., and Amini., A.M. 2002. Quantitative land suitability evaluation in north Baron Region (esfahan province) for wheat, barley, maize and rice, *Journal of Agricultural and Natural resource research and Technology*, 3(6): 105-118. (In Persian).
5. Ayoubi SH.A., and Jalalian A. 2010. Land Evaluation (Agricultural and Natural Resources Secend Edition), Isfahan University of Technology Publication Center, Isfahan, Iran, 385p. (In Persian).
6. Azadeh A., Saberi M., Gitiforouz Z., and Saberi Z. 2009. A Hybrid simulation adaptive network based fuzzy inference system for improvement of electricity consumption estimation. *Expert System Applied Journal*, 36(8): 11108-11117.
7. Bremner J.M. 1965. Inorganic forms of nitrogen. In: Black (Ed.), *Methods of soil analysis*. Part 2, Monograph NO, 9: 1179-1232. Am. Soc. Agron., Madison, WI.
8. De la Rosa D., Moreno JA., Garcia LV., and Almorza, J. 1992. Microleis: A microcomputer-based Mediterranean land evaluation information system. *Soil Use and Management* 8: 89-96.
9. Demuth H., and Beale, M. 2000. *Neural Network Toolbox Users Guideline*. By The Math Works, Inc, Version 4, 840pp.
10. Dep K., and Beyer G. 2001. Self-adaptive Genetic Algorithms with Simulated Binary Cross over. *Evolution Computer Journal*, 9:197-221.
11. Eftekhari Ardebili, M. 2004. Artificial Neural Network Modelling. MSc Thesis of AL Zahra university of Tehran, Tehran.
12. Etedali S., Givi J., and Nouri M.R. 2012. Comparison Between Land Production Potential Prediction for Maize, Using FAO and Wageningen Models and Assessment of Management Level for Its Cultivation Around Shahrekord City. *Journal of Water and soil*, 26 (4): 873-855. (in Persian).
13. Gee G. W., and Bauder J. W. 1986. Particle size analysis. p. 383-411. In: A. Klute (Ed), *Methods of soil analysis*. Part I. Physical and mineralogical methods, 2nd ed., Agronomy Monograph. No: 9. ASA and SSSA. Madison, WI.
14. Halder J.C. 2013. Land suitability assessment for crop cultivation by using remote sensing and GIS. *Journal of Geography and Geology*, 5(3): 65-78.
15. Jalalian A., Rostaminia M., Ayoubi SH.A., and Amini Am. Qualitative, Quantitative and economic land suitability evaluation for wheat, maize and sesame in Mehran plain. *Journal of Agricultural and Natural resource research and Technology*, 4(3): 393-403. (In Persian).
16. Kim S., and Kim S.H. 2008. Neural Networks and genetic algorithm approach for nonlinear evaporation and evapotranspiration modeling. *Journal of Hydrology*, 351: 299-317.
17. Koza J.R. 2008. <http://www.genetic-programming.com>, The home page of john R. Koza at Genetic Programming Lnc.

18. Liu J., and Goering C.E. 1999. Neural network for setting target corn yields. ASAE paper 99-3040, Toronto, Ontario, Canada, July 18-21.
19. Manhaj, M.B. 2005. Fundamental of Artificial Neural Network. Amir kabir University of Technology Center, 712p.
20. McLean E.O. 1982. Soil pH and Lime requirement. Pp. 199-224. In: Page AL, Miller RH and Keeney DR (eds). Methods of Soil Analysis. Part 2. Chemical and Micromorphological Properties. 2nd ed. Agron, Monogr. 9. ASA and SSSA, Madison, WI.
21. Mohammadi J. 2006. Pedometrics: Classical Statistics. Pelk publication. (in Persian).
22. Mohammadi J. 2006. Pedometrics: Fuzzy Systems Theory. Pelk publication. (in Persian).
23. Nelson R.E. 1982. Carbonate and gypsum. Pp. 181-197. In: Page AL, Miller RH and Keeney DR (eds). Methods of Soil Analysis. Part 2. Chemical and Microbiological Methods. 2nd ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI.
24. Nelson R.E., and Sommers L. 1982. Total carbon, organic carbon and organic matter. Pp. 532-581. In: Page AL, Miller RH and Keeney DR (eds). Methods of Soil Analysis. Part 2. Chemical and Microbiological Methodes. 2nd ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI.
25. Newhall F., and Berdanier C.R. 1996. Calculation of soil moisture regimes from the climatic record. Natural Resources Conversations Service, Soil Survey Investigation. Report No. 46, 13p.
26. Norouzi M., Ayoubi SH.A., Jalalian A., Khademi H. and Dehghani, A. 2010. Predicting rainfed wheat quality and quantity by artificial neural network using terrain and soil characteristics. Acta Agriculturae Scandinavica Section B-Soil and Plant Science, 60(4): 341-352.
27. Olsen S.R., Cole C.V., Watanable, F.S., and Dean, L.A. 1954. Estimation of available phosphorus in soil by extraction with sodium bicarbonate. USDA. Circ. 939. U. S. Government. Print. Office, Washington, DC.
28. Sayegh A.H., Khan P., and Ryan, J. 1978. Factors affecting gypsum and cation exchange capacity determination in gypsiferous soils. Soil Science Journal, 125: 294-300.
29. Schoeneberger P.J., Wysocki D.A., Benham E.C. and Broderson, W.D. 2006. Field Book for Describing and Sampling Soils. Natural Resources Conservation Service, USDA, National Soil Survey Center, Lincoln, NE, 280p
30. Servati M, Jafarzadeh, A.A., Ghorbani M.A., Shahbazi F. and Davatgar N. 2014. Comparison of the FAO and Albero Models, in Prediction of Irrigated Wheat Production Potentials in the Khajeh region. Water and Soil Science Journal, 24: 1-14. (in Persian).
31. Sette S., and Boullart L. 2001. Genetic programming: principles and applications. Engineering Applications of Artificial Intelligence, 14: 727-736.
32. Shite N, Khaje. Geological survey and Mineral Exploration of Iran. 2006. Geology Map of Iran, 1:100000series.
33. Sohrabi vafa H. 2012. Forecasting of energy demand in Iran by soft computation technic. Msc thesis of shahid abbaspour energy and water Industry university.
34. Soil Survey Field and Labra Methods Manuals. 2014. Soil Survey Investigations Report No.51. Version.2. Issued 2014.
35. Sudduth K.A., Drummond S.T., Birrell S.J., and Kitchen, N.R. 1996. Analysis of spatial factors influencing crop yield, in Proc. 3rd Int. Conf. On Precision Agriculture, P.C. Robert et al. (ed.), pp. 129-140.
36. Sys C., Van Ranset E., and Debaveye J. 1991a. Land Evaluation, Part I, Principle in Land Evaluation and Crop Production Calculation, International Training Center for Post Graduate Soil Scientists, Ghent University, Ghent., Belgium.
37. Sys C., Van Ranset E., and Debaveye J. 1991b. Land Evaluation, Part II, Methods in Land Evaluation. International Training Center for Post Graduate Soil Scientists, Ghent University, Ghent, Belgium.
38. Taghizadeh Mehrjerdi R., Seyedjalali S.A., and Sarmadian F. 2016. Prediction of Corn Spatial yield by soil digital mapping in Gotend region (Khuzestan Province, IRAN). Jornal of plant production, 19 (4): 70-96.
39. Yang Y., and Cai Y. 2000. Sustainable evaluation on rural resources, environment and development of China – the SEEA method and its applications. Acta Geographica Sinica, 55 (5): 596-606.



Prediction of Irrigated Wheat Yield by using Hybrid Algorithm Methods of Artificial Neural Networks and Genetic Algorithm

A. Barikloo¹- P. Alamdari*²- K. Moravej³ – M. Servati⁴

Received: 16-06-2016

Accepted: 08-03-2017

Introduction: In recent decades, the most important issue for agricultural activities is maximizing the productions. Today, wheat is grown on more lands than any other commercial crops and continues to be the most important food grain source for humans. Sustainable agriculture is a scientific activity based on ecological principles with focus on achieving sustainable production. It requires a full understanding of the relationships between crop production with soil and land characteristics. Furthermore, one of the objectives of sustainable agriculture is enhancing the agricultural production efficiency through applying proper management, which requires a deep understanding of relationships between production rate, soil and environment characteristics. Hence, the first step in this process is finding appropriate methods which are able to determine the correct relationships between measured characteristics of soil and environment with performance rate. The aim of this study was evaluating the performance of neuro-genetic hybrid model in predicting wheat yield by using land characteristics in the west of Heris City.

Materials and Methods: The study area was located in the northwest of east Azarbaijan province, Heris region. In this study, 80 soil profiles were surveyed in irrigated wheat farms and soil samples were taken from each genetic horizon for physical and chemical analyses. In this region, soil moisture and temperature regimes are Aridic border to Xeric and Mesic, respectively. The soils were classified as Entisols and Aridisols. We used 1×1 m woody square plots in each profile to determine the amounts of yield. Because of nonlinear trend of yield, a nonlinear algorithm hybrid technique (neural-genetics) was used for modeling. At first step, the average weight of soil characteristics (from depth of 100 cm) and landscape parameters of selected profiles were measured for modeling according to the annual growing season of wheat. Then, land components and wheat yield were considered as inputs and output of model, respectively. For this reason, genetic algorithm was investigated to train neural network. Finally, estimated wheat yield was obtained using input data. Root mean square error (RMSE) and Coefficient of determination (r^2), Nash-Sutcliffe Coefficient (NES) indices were used for assessing the method performance.

Results and Discussion: The sensitivity analysis of model showed that soil and land parameters such as total nitrogen, available phosphorus, slope percentage, content of gravel, soil reaction and organic matter percentage played an important role in determining wheat yield in the studied area. The soil organic matter and total nitrogen had the highest and lowest correlation with wheat yield quantity and quality, respectively, indicating the total nitrogen was the most important soil property for determination of wheat yield in our studied area. We found that network learning process based on genetic algorithms in the learning process had lower error. The findings showed that beside of confirming the desired results in the case of using sigmoid activation function in the hidden layer and linear activation function in the output layer of all neural networks, it is demonstrated that the proposed hybrid technique had much better results. These findings also confirm better prediction ability of neural network based on error back propagation algorithm or Levenberg-Marquardt training algorithm compared to other types of neural network confirms.

Conclusion: Using nonlinear techniques in modeling and forecasting wheat yield due to its nonlinear trend and influencing variables is inevitable. Recently, genetic algorithms and neural network techniques is considered as the most important tools to model nonlinear and complex processes. Despite the advantages of these techniques there are a lot of weaknesses. Imposing specific conditioned form by researchers in the techniques of genetic algorithms and stopping neural network learning at the optimal points are the main weaknesses of these

1, 2 and 3- M.Sc. Graduated and Assistant Professors of Soil Science Department, Faculty of Agriculture, University of Zanzan, Respectively

(*-Corresponding Author Email: p_alamdari@znu.ac.ir)

4- Assistant Professor of Shahid Bakeri High Education Center of Miandoab, Urmia University

techniques, while searching for global optimal point and not imposing a specific functional forms are the robustness of genetic algorithm techniques and neural networks, respectively. Results of this study indicated that the proposed hybrid technique had much better results. Correlation coefficient (0.87) and average deviation square error (473.5) were high and low, respectively. It can be concluded that the surveyed soil properties have very strong relationship with the yield. Implementation of appropriate land management practices is thus necessary for improving soil and land characteristics to maintain high yield, preventing land degradation and preserving it for future generations required for sustainable development.

Keywords: Hard and Readily measured properties, Management, Modelling, Sustainable agriculture