

تعیین فاکتورهای موثر بر آتش‌سوزی جنگل با استفاده از ترکیب مدل رگرسیون اسپلاین

تطبیقی چندمتغیره و الگوریتم ژنتیک

مطالعه موردی: جنگل گلستان

پرهام پهلوانی^۱، استادیار دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی دانشگاه

تهران، ایران

امین راعی، دانشجوی دکترای مهندسی سیستم‌های اطلاعات مکانی، دانشکده مهندسی نقشه‌برداری و اطلاعات

مکانی، پردیس دانشکده‌های فنی دانشگاه تهران، ایران

بهناز بیگدلی، استادیار دانشکده مهندسی عمران، دانشگاه صنعتی شاهرود، ایران

پذیرش نهایی: ۱۳۹۸/۱۰/۰۷

دریافت مقاله: ۱۳۹۷/۰۵/۰۸

چکیده

امروزه شناسایی فاکتورهای موثر بر آتش‌سوزی جنگل‌ها از اهمیت بسیار بالایی برخوردار است زیرا سالانه مساحت زیادی از جنگل‌های جهان بر اثر آتش‌سوزی نابود می‌شوند و تکرار این اتفاق در بلندمدت می‌تواند خسارات جبران‌ناپذیری بر زمین و ساکنین آن وارد کند. با شناسایی این فاکتورها می‌توانیم زمان‌ها و نقاط دارای ریسک بالای آتش‌سوزی را شناسایی نمائیم و با وضع قوانین و سیاست‌های مدیریتی کارآمد، آموزش به مردم و نظارت بیشتر در جهت مقابله با عوامل محرک آتش برآییم. در این تحقیق سعی شده است فاکتورهای موثر بر آتش‌سوزی‌های جنگل گلستان شناسایی شود و برای این منظور از سه روش رگرسیون خطی چندگانه، رگرسیون لجستیک و رگرسیون اسپلاین تطبیقی چندمتغیره در ترکیب با الگوریتم ژنتیک استفاده شد. نتایج این تحقیق نشان داد که هر دو دسته فاکتورهای بیوفیزیکی و انسانی در آتش‌سوزی‌های منطقه مورد مطالعه دارای تأثیر هستند. از این میان تنها فاکتورهای حداقل دما و حداکثر سرعت باد در هر سه حالت موثر شناخته شدند. روش رگرسیون اسپلاین تطبیقی چندمتغیره در مقایسه با دو روش دیگر عملکرد بهتری از خود نشان داد. مقدار RMSE نرمال شده این سه روش برابر ۰/۴۲۹۱ برای رگرسیون خطی چندگانه، ۰/۹۴۱۶ برای رگرسیون لجستیک و ۰/۱۷۵۷ برای رگرسیون اسپلاین تطبیقی چندمتغیره و مقدار R^2 آن‌ها نیز به ترتیب برابر ۰/۹۸۶۲، ۰/۹۹۱۲ و ۰/۹۸۸۶ به دست آمد.

واژه‌های کلیدی: آتش‌سوزی جنگل، رگرسیون اسپلاین تطبیقی چندمتغیره، رگرسیون خطی چندگانه، رگرسیون لجستیک،

الگوریتم ژنتیک

مقدمه

امروزه، آتش‌سوزی جنگل‌ها از جمله مهم‌ترین و زیان‌بارترین بلایای طبیعی جهان به شمار می‌آید و تعیین فاکتورهای موثر بر ایجاد و توسعه آتش به منظور تشخیص به‌موقع آن و واکنش بدان ضروری به نظر می‌رسد. تاکنون مطالعات زیادی در این زمینه انجام شده است. در بیشتر موارد انسان عامل وقوع آتش‌سوزی بوده است (Martinez-Fernandez et al., ۲۰۱۳) اما فاکتورهای انسانی موثر بر آتش‌سوزی جنگل‌ها هنوز به طور کامل شناسایی نشده است.

بیشتر فاکتورهای موثر بر آتش‌سوزی جنگل که در تحقیقات گذشته مورد مطالعه قرار گرفته‌اند در دو گروه اصلی فاکتورهای بیولوژیکی- فیزیکی یا به طور خلاصه بیوفیزیکی و فاکتورهای اقتصادی- اجتماعی یا انسانی قرار می‌گیرند. در بسیاری از موارد فاکتورهای بیوفیزیکی مانند نرخ بارگذاری سوخت تنها فاکتورهایی بودند که در مدیریت بحران در نظر گرفته شدند (Dickson et al., ۲۰۰۶). این در حالی است که انسان با ایجاد یا خاموش کردن حریق تأثیر مستقیمی در روند آتش‌سوزی جنگل‌ها دارد. طبق تحقیقات انجام شده توسط موریتز و همکاران (Moritz et al., ۲۰۰۴) نیز، فاکتورهای بیوفیزیکی به تنهایی توانایی تعیین الگوی آتش‌سوزی جنگل‌ها را ندارند و برای این کار باید با فاکتورهای انسانی ترکیب شوند.

اگرچه به طور کلی آتش‌سوزی رفتاری تصادفی دارد (Yang et al., ۲۰۰۷) اما کانتریمن (Contryman, ۱۹۷۲) نشان داد که می‌توان روند اصلی آن را به صورت تابعی از سه نوع متغیر در نظر گرفت، شامل (الف) شرایط جوی منطقه مانند دما، رطوبت هوا، میزان بارندگی، سرعت وزش باد و جهت عمده آن؛ (ب) ویژگی‌های سوختی منطقه مانند نوع مواد سوختنی، نرخ بارگذاری سوخت، رطوبت مواد سوختنی و اشتعال‌پذیری آن‌ها؛ و (ج) توپوگرافی منطقه مانند ارتفاع، شیب و جهت شیب که هر سه دسته جزء فاکتورهای بیوفیزیکی به شمار می‌آیند. به همین دلیل بخش عمده‌ای از تحقیقات انجام شده بر فاکتورهای بیوفیزیکی متمرکز بوده و کمتر به فاکتورهای انسانی توجه نموده‌اند. از جمله تحقیقاتی که در زمینه مدل‌سازی توسعه آتش در شرایط بیوفیزیکی مختلف انجام شده (Burgan & Rothermel, ۱۹۸۴; Forestry Canada, ۱۹۹۲) و مدل‌های زیادی که جهت شناسایی مناطق مستعد آتش‌سوزی با استفاده از فاکتورهای بیوفیزیکی ارائه شده‌اند (Larjavaara et al., ۲۰۰۵; Wotton & Martell, ۲۰۰۵).

تاکنون مدل‌های متنوعی جهت شناسایی فاکتورهای موثر بر آتش‌سوزی جنگل مورد استفاده قرار گرفته است و در نتایج آن‌ها، هر دو دسته فاکتورهای بیوفیزیکی و انسانی حضور دارند. به عنوان مثال طبق نتایج رودریگز-ترخو و همکاران (Rodriguez-Trejo et al., ۲۰۰۸) در مکزیک حداکثر سرعت باد، تعداد تولیدکنندگان آتش و سطح تحصیلات آن‌ها بیشترین اثر را در آتش‌سوزی جنگل‌ها داشته است. نتایج رومن-کوئستا و همکاران (Roman-Cuesta et al., ۲۰۰۶) حاکی از وجود ارتباطی معنادار بین آتش‌سوزی جنگل‌ها، میزان تراکم جاده‌ها و نوع کاربری کشاورزی در منطقه مورد مطالعه آن‌ها می‌باشد. یانگ و همکاران (Yang et al., ۲۰۰۷) با استفاده از آنالیزهای آماری نشان دادند که در منطقه مورد مطالعه آن‌ها احتمال وقوع آتش‌سوزی در محل‌های در دسترس انسان‌ها، نقاط واقع در شیب‌های متوسط و ارتفاعات بالاتر جنگل و در جنگل‌های کاج و بلوط بیشتر است. سیفارد و همکاران (Syphard et al., ۲۰۰۸) الگوی مکانی آتش‌سوزی‌های جنگل‌ها را با استفاده از مدل‌های رگرسیون لجستیک چندگانه^۱ و رگرسیون پواسون چندگانه^۲ مدل‌سازی

^۱ Multiple logistic regression

^۲ Poisson multiple regression

کردند و نتیجه گرفتند که در منطقه مورد مطالعه آن‌ها تعداد آتش‌سوزی‌ها بیشتر تحت تأثیر فاکتورهای انسانی و گستره آن‌ها تحت تأثیر فاکتورهای بیوفیزیکی است. رومرو-کالسرادا و همکاران (Romero-Calcerrada et al., ۲۰۰۸) با استفاده از مدل وزن‌های نشانگر^۱ نشان دادند که میزان دسترسی انسان‌ها به جنگل بیشترین تأثیر را در آتش‌سوزی‌های مرکز اسپانیا دارد.

هدف ما در تحقیق پیش رو تعیین فاکتورهای موثر بر آتش‌سوزی‌های بخشی از جنگل گلستان با استفاده از مدل‌های رگرسیون خطی چندگانه^۲ (MLR)، رگرسیون لجستیک^۳ (LR) و رگرسیون اسپلاین تطبیقی چندمتغیره^۴ (MARS) و مقایسه آن‌ها با یکدیگر بوده است. طبق آمار آتش‌سوزی‌های منطقه در سال‌های اخیر بیشتر آن‌ها دارای منشأ انسانی بوده‌اند؛ بنابراین در این تحقیق هر دو دسته فاکتورهای بیوفیزیکی و انسانی را در نظر گرفتیم و فرض ما بر این است که فاکتورهای انسانی دارای اثرگذاری بیشتری هستند. از آنجایی که ترکیب‌های متفاوت فاکتورها به پاسخ‌های متفاوتی می‌انجامد بنابراین لازم است ترکیب بهینه‌ای از فاکتورها را بیابیم که بدین منظور از الگوریتم ژنتیک استفاده نمودیم. در ادامه تحقیق حاضر، به معرفی منطقه مورد مطالعه، داده‌های مورد استفاده، روش رگرسیون خطی چندگانه، رگرسیون لجستیک، رگرسیون اسپلاین تطبیقی چندمتغیره و الگوریتم ژنتیک مورد استفاده جهت یافتن بهترین ترکیب فاکتورها می‌پردازیم. سپس نتایج به دست آمده ارائه و پیرامون آن‌ها بحث می‌شود و در نهایت به نتیجه‌گیری از این تحقیق می‌پردازیم.

داده‌ها و روش کار

حدود ۱۸٪ از سطح استان گلستان پوشیده از جنگل است و مناطق جنگلی این استان از مناطق توریستی کشور به شمار می‌آید. همچنین جاده‌های ارتباطی متعددی از جنگل‌های آن عبور می‌کند و این امر دسترسی مردم به جنگل را تسهیل می‌کند. طبق آمار ثبت شده اکثر آتش‌سوزی‌های جنگلی استان در مجاورت این جاده‌ها رخ داده‌اند. منطقه مورد مطالعه این تحقیق (شکل ۱)، در محدوده $37^{\circ}25' - 36^{\circ}53'$ شمالی و $55^{\circ}50' - 55^{\circ}5'$ شرقی واقع شده است و مساحتی در حدود $3719/5$ کیلومتر مربع دارد. دلیل انتخاب این منطقه این بوده است که بیشتر آتش‌سوزی‌های اتفاق افتاده اخیر در استان گلستان در این محدوده بوده است از جمله آتش‌سوزی بزرگی که در ۲۶ آذر سال ۸۹ رخ داد و از اطلاعات آن در این تحقیق استفاده شده است.

بخشی از داده‌های مورد نیاز این تحقیق شامل مناطق واقعی سوخته شده در آتش‌سوزی تاریخ ۲۶ آذر ۱۳۸۹ (شکل ۲) و داده‌های دیگری از قبیل مدل رقومی ارتفاعی^۵ (DEM) منطقه، شبکه راه‌ها، رودخانه‌ها، جنس خاک، کاربری اراضی و مناطق مسکونی (شکل ۳) از سازمان نقشه‌برداری کل کشور و سازمان منابع طبیعی استان گلستان تهیه شده است. همچنین داده‌های ۵ ایستگاه هواشناسی سینوپتیک نزدیک به منطقه مورد مطالعه (شکل ۴) شامل حداکثر، حداقل و متوسط دما، میزان بارندگی، جهت عمده وزش باد و حداکثر سرعت باد در آذر ماه سال ۸۹ به همراه مختصات جغرافیایی این ایستگاه‌ها از طریق سازمان ملی هواشناسی کشور به دست آمده است. چون مقدار این پارامترها در منطقه مورد

^۱ Weights of evidence

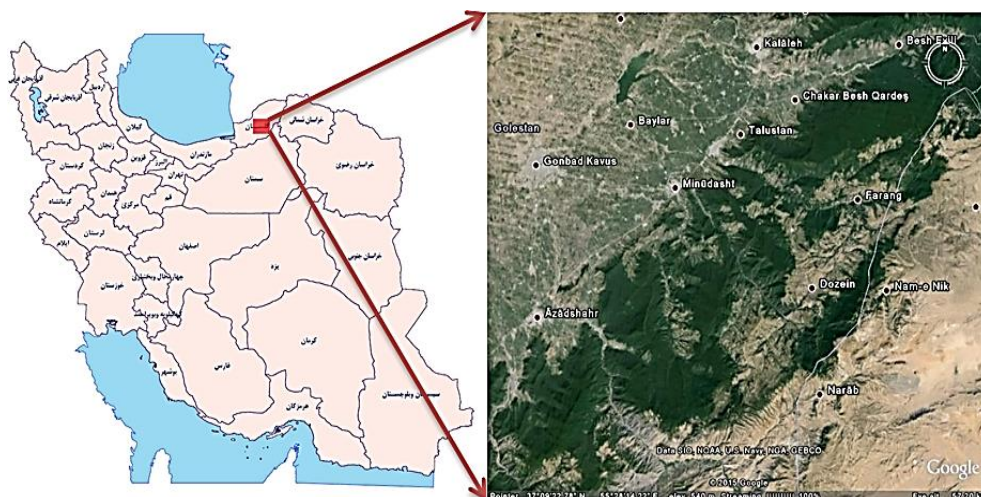
^۲ Multiple linear regression

^۳ Logistic regression

^۴ Multivariate adaptive regression spline

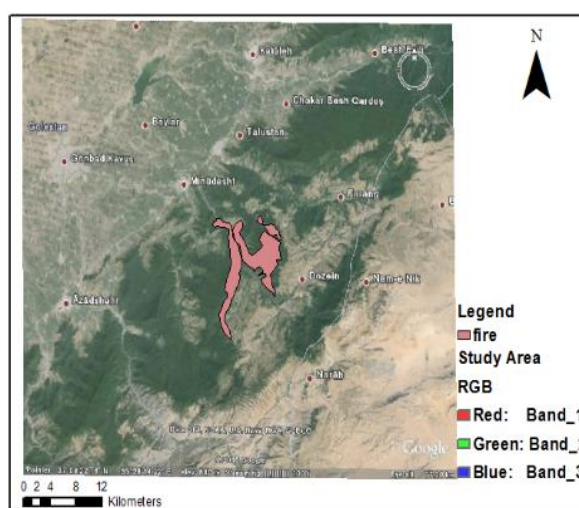
^۵ Digital Elevation Model

مطالعه در دست نبوده و اساساً قابلیت مشاهده این پارامترها در کل منطقه وجود ندارد، برای به دست آوردن مقادیر آنها در منطقه مورد مطالعه از درون‌یابی به روش کریجینگ عمومی^۱ با مدل نمایی و قدرت تفکیک ۳۰ متر استفاده نمودیم (شکل ۵).



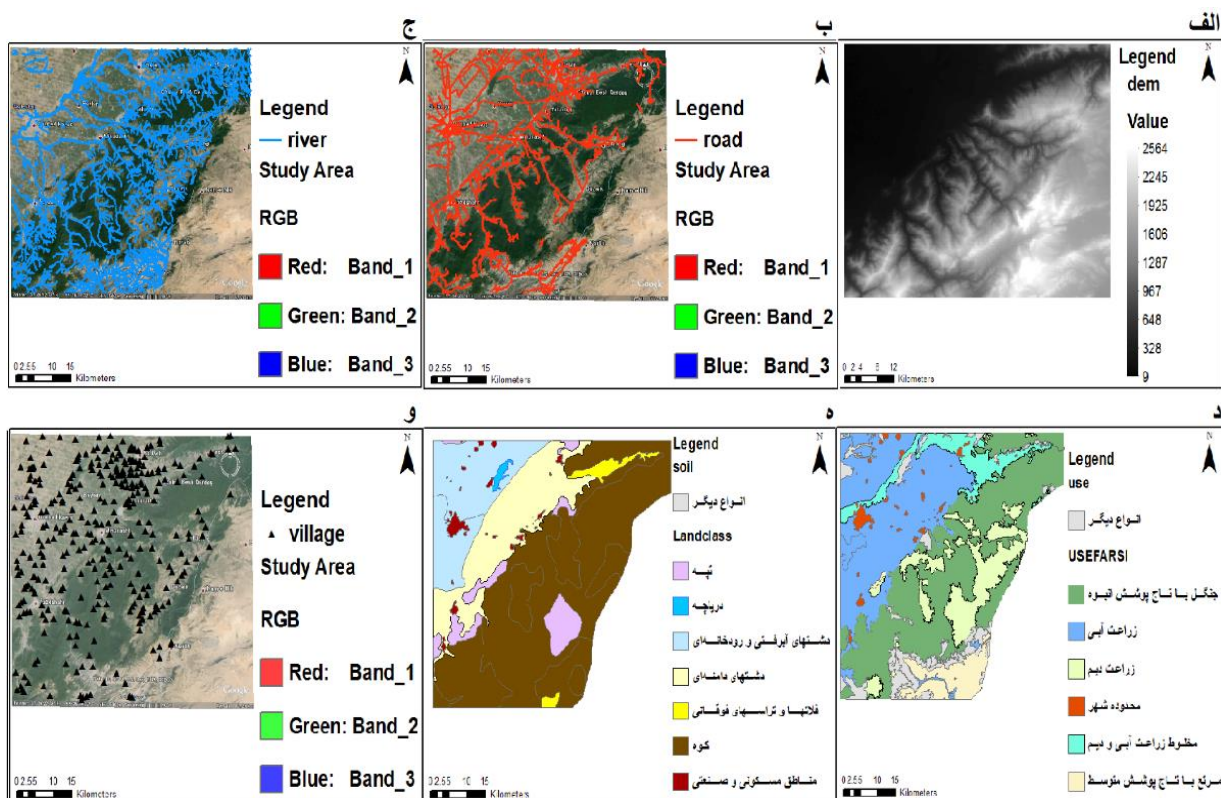
شکل ۱: منطقه مورد مطالعه

لایه‌های اطلاعاتی جنس خاک و کاربری اراضی در مقیاس ۱:۱۰۰۰۰۰ و شبکه راه‌ها و رودخانه‌ها در مقیاس ۱:۵۰۰۰ تهیه شده‌اند و مربوط به سال ۸۵ می‌باشند. مدل رقومی ارتفاعی منطقه از نقشه‌های توپوگرافی ۱:۲۵۰۰۰ سازمان نقشه برداری کل کشور و با قدرت تفکیک ۳۰ متر تهیه شده و با استفاده از آن در نرم‌افزار ArcGIS لایه‌های اطلاعاتی شیب و جهت شیب با قدرت تفکیک ۳۰ متر تولید شده است (شکل ۶). لایه‌های اطلاعاتی راه‌ها، رودخانه‌ها و مناطق مسکونی دارای فرمت برداری بودند بنابراین با استفاده از آنالیز فاصله اقلیدسی، رستریابی با قدرت تفکیک ۳۰ متر تولید گشت که هر سلول آن نشان‌دهنده فاصله از نزدیک‌ترین جاده، رودخانه یا منطقه مسکونی است (شکل ۷).

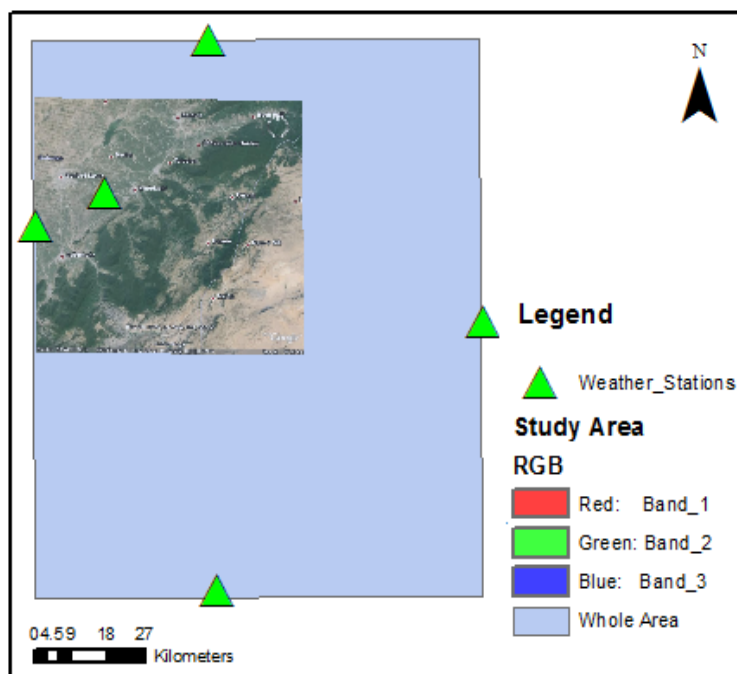


شکل ۲: آتش‌سوزی اتفاق افتاده در منطقه مورد مطالعه در تاریخ ۲۶ آبان ۱۳۸۹

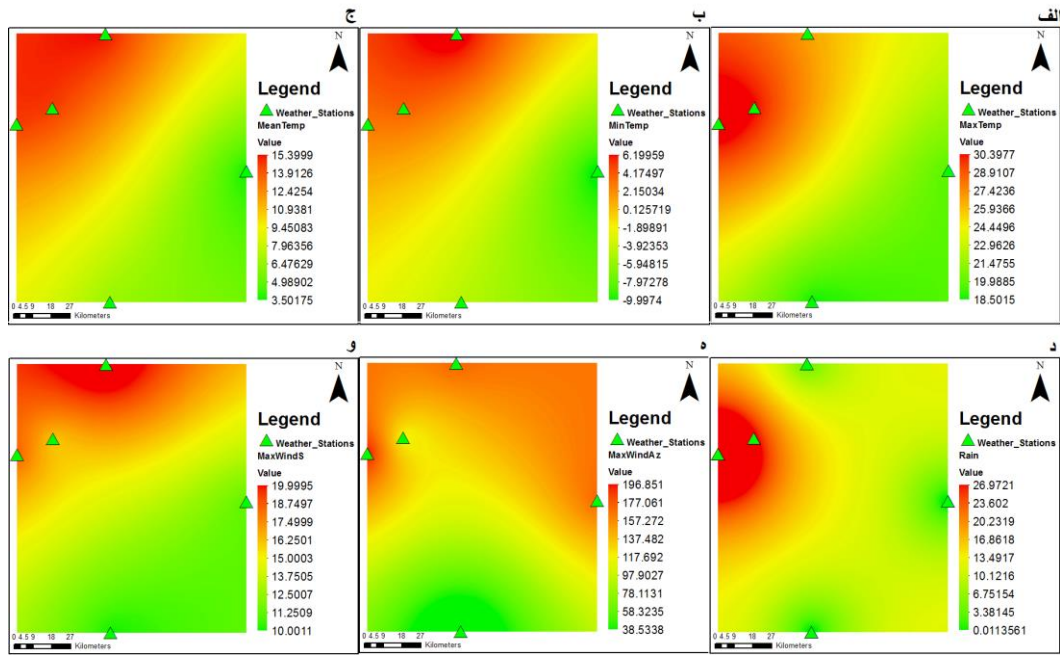
^۱ Ordinary kriging



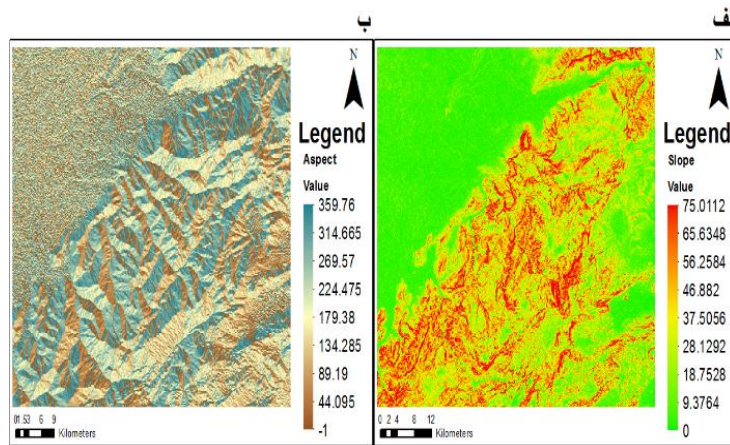
شکل ۳: لایه‌های اطلاعاتی (الف) مدل رقومی ارتفاعی منطقه، (ب) شبکه راه‌ها، (ج) رودخانه‌ها، (د) کاربری اراضی، (ه) جنس خاک و (و) مناطق مسکونی



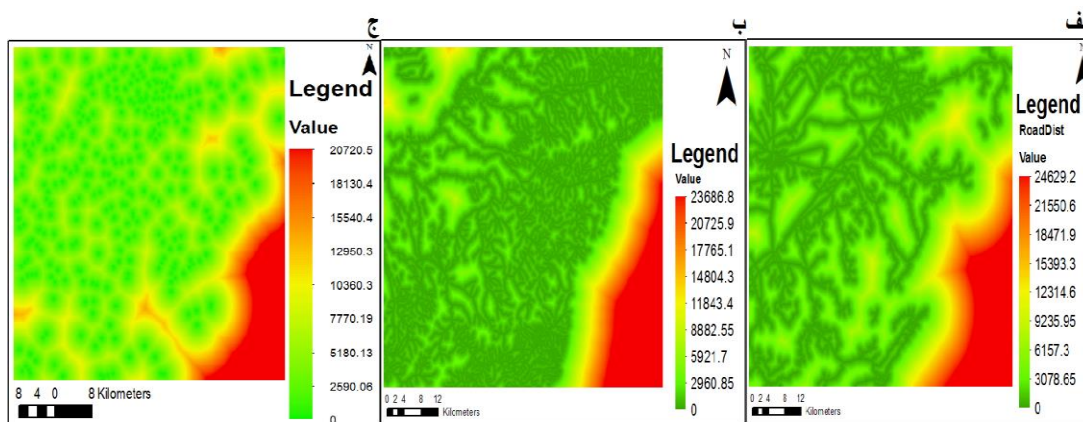
شکل ۴: ایستگاه‌های هواشناسی سینوپتیک نزدیک به منطقه مورد مطالعه



شکل ۵: رسترهای تولید شده از درون‌یابی (الف) حداکثر دما (°C)، (ب) حداقل دما (°C)، (ج) متوسط دما (°C)، (د) حداکثر سرعت باد (M/S)، (ه) جهت باد غالب و (و) میزان بارندگی (MM) بین ایستگاه‌های هواشناسی نزدیک به منطقه مورد مطالعه

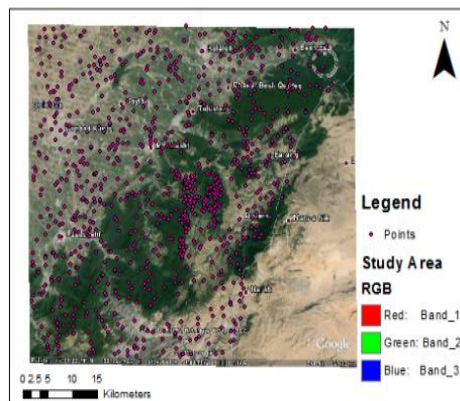


شکل ۶: لایه‌های اطلاعاتی (الف) شیب، (ب) جهت شیب



شکل ۷: لایه‌های اطلاعاتی فاصله از (الف) مناطق مسکونی (M)، (ب) رودخانه‌ها (M) و (ج) جاده‌ها (M)

در نهایت ۱۰۰۰ نقطه با توزیع تصادفی یکنواخت البته با تراکم بیشتر در محدوده آتش در منطقه مورد مطالعه تولید شد (شکل ۸). سپس، مقادیر همه لایه‌های اطلاعاتی موجود در آن‌ها به دست آمد. از این تعداد ۶۰۰ نقطه را به عنوان نقاط مشاهداتی و ۴۰۰ نقطه را به عنوان نقاط چک انتخاب نمودیم.



شکل ۸: نقاط تصادفی تولید شده در منطقه مورد مطالعه

نخستین مدل رگرسیون مورد استفاده در این تحقیق رگرسیون خطی چندگانه است. این مدل در واقع تعمیم‌یافته رگرسیون خطی ساده می‌باشد که در آن ارتباط خطی یک متغیر به چند متغیر مستقل مدل‌سازی می‌شود و رابطه کلی آن به صورت زیر است:

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n + e, \quad (1)$$

که در آن y متغیر وابسته، x_1, \dots, x_n متغیرهای مستقل، a_0, \dots, a_n ضرایب رگرسیون و e باقیمانده مدل است. هدف از رگرسیون نیز تعیین ضرایب رگرسیون است که در واقع تأثیر هر کدام از متغیرهای مستقل را در متغیر وابسته نشان می‌دهند. رابطه (۱) را می‌توان به صورت زیر نیز بازنویسی کرد:

$$y = X_{1 \times (n+1)} A_{(n+1) \times 1} + e, \quad (2)$$

$$X = [1 \ x_1 \ x_2 \ \dots \ x_n], \quad (3)$$

$$A = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \quad (4)$$

با استفاده از سرشکنی به روش کمترین مربعات^۱، ضرایب مجهول از رابطه زیر به دست می‌آید:

$$\hat{A} = (X^T X)^{-1} X^T y, \quad (5)$$

که در آن \hat{A} برآوردی از ضرایب رگرسیون است.

دومین مدل رگرسیون مورد استفاده در این تحقیق مدل رگرسیون لجستیک است که یک مدل ریاضی غیرخطی برای تعیین رابطه بین یک متغیر وابسته دودویی و چند متغیر مستقل است (Kong et al., ۲۰۱۶). در این تحقیق چنانچه از

^۱ Least squares

مقادیر ۱ و ۰ به ترتیب برای نشان دادن نقاط آتش و غیرآتش استفاده شود و همچنین فرض شود که احتمال آتش بودن یک نقطه از تابع لجستیک زیر که توسط کلینباوم و کلین (Kleinbaum & Klein, ۱۹۹۴) ارائه شده است، پیروی کند:

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

آنگاه می‌توان احتمال آتش بودن یک نقطه را از رابطه زیر محاسبه کرد:

$$P(y = 1 | x_1, x_2, \dots, x_n) = \frac{1}{1+e^{-(\alpha+\sum_{i=1}^k \beta_i x_i)}}, \quad (7)$$

که در آن $P(y=1 | x_1, x_2, \dots, x_n)$ احتمال شرطی آتش بودن متغیر وابسته به ازای متغیرهای مستقل x_1 تا x_n است و α و β_i ضرایب رگرسیون هستند. برای محاسبه ضرایب مجهول این مدل در مواردی که تعداد پارامترها در مقایسه با مشاهدات ناچیز است، از برآورد به روش ماکزیمم شباهت غیرمشروط^۱ استفاده می‌شود که رابطه آن به صورت زیر است (Kleinbaum & Klein, ۱۹۹۴):

$$L_U = \frac{\prod_{i=1}^n \exp(\alpha + \sum_{i=1}^k \beta_i x_{il})}{\prod_{i=1}^n [1 + \exp(\alpha + \sum_{i=1}^k \beta_i x_{il})]}, \quad (8)$$

که در آن ضرایب α و β_i به گونه‌ای تغییر می‌کنند که مقدار L_U ماکزیمم گردد. سومین مدل رگرسیون مورد استفاده در این تحقیق، مدل رگرسیون اسپلاین تطبیقی چندمتغیره است که در سال ۱۹۹۱ توسط فریدمن (Friedman, ۱۹۹۱) ارائه شد. MARS نیاز به هیچ‌گونه پیش‌فرضی در مورد نوع رابطه متغیرهای وابسته و مستقل ندارد؛ به همین دلیل در مقایسه با بسیاری از مدل‌های دیگر از توانایی بالایی در تعیین رابطه بین متغیرها برخوردار است و می‌تواند روابط پیچیده و غیرخطی بین آن‌ها را مدل کند (West et al., ۲۰۱۸). رابطه کلی MARS به صورت زیر است (Friedman, ۱۹۹۱):

$$\hat{y} = \hat{f}(X) = a_0 + \sum_{m=1}^M a_m BF_m(X), \quad (9)$$

که در آن \hat{y} برآوردی از متغیر وابسته و به صورت تابعی از متغیرهای مستقل X جمله ثابت، a_0 تعداد توابع پایه غیرثابت یا تعداد گره‌ها، a_m ضریب مربوط به تابع پایه m ام و $BF_m(X)$ تابع پایه m ام است که از رابطه زیر به دست می‌آید (Friedman, ۱۹۹۱):

$$BF_m(X) = \prod_{i=1}^{K_m} [S_{i,m}(X_{v(i,m)} - t_{i,m})]_+^q, \quad (10)$$

که در آن K_m درجه تعامل بین متغیرها در تابع پایه m ام، $S_{i,m}$ برابر با ± 1 ، $X_{v(i,m)}$ متغیر مستقل v ام، $t_{i,m}$ محل گره هر یک از متغیرهای مستقل حاضر در تابع پایه m ام، q درجه تابع پایه و اندیس $+$ به معنای انتخاب بخش مثبت عبارت $[S_{i,m}(X_{v(i,m)} - t_{i,m})]_+^q$ است یعنی:

$$[S_{i,m}(X_{v(i,m)} - t_{i,m})]_+^q = \max(0, [S_{i,m}(X_{v(i,m)} - t_{i,m})]^q), \quad (11)$$

روش اعتبارسنجی متقابل تعمیم‌یافته^۲ (GCV) که از آن برای جلوگیری از افزونگی در MARS استفاده می‌شود، به صورت زیر است (Craven & Wahba, ۱۹۷۹):

$$GCV(M) = \frac{1}{N} \sum_{i=1}^N [y_i - \hat{f}(X_i)]^2 / \left[1 - \left(\frac{C(M)}{N}\right)\right]^2, \quad (12)$$

که در آن M تعداد توابع پایه غیرثابت، N تعداد داده‌ها، y_i مقدار واقعی متغیر وابسته، $\hat{f}(X_i)$ مقدار برآوردشده متغیر وابسته و $C(M)$ معیار هزینه مدل و تعداد موثر درجه آزادی می‌باشد. در مرحله حرکت رو به عقب پس از حذف هر تابع

^۱ Unconditional maximum likelihood

^۲ Generalized Cross Validation

پایه مقدار GCV محاسبه می‌گردد و هر تابعی که با حذف آن مقدار GCV کاهش یابد، از مدل کنار گذاشته می‌شود تا در نهایت مدلی با مقدار GCV بهینه داشته باشیم. فریدمن (Friedman, ۱۹۹۱) رابطه زیر را برای تعیین $C(M)$ پیشنهاد داد:

$$C(M) = \text{trace}(B(B^T B)^{-1} B^T) + 1 + d.M, \quad (13)$$

که در آن B یک ماتریس $M*N$ شامل مقدار توابع پایه به ازای داده‌های ورودی و d هزینه به ازای هر تابع پایه است که معمولاً عددی بین ۲ تا ۴ در نظر گرفته می‌شود. هر چقدر d بزرگتر انتخاب شود تعداد گره‌ها کمتر و مدل نهایی هموارتر خواهد بود. به همین دلیل آن را به عنوان پارامتر هموارسازی نیز در نظر می‌گیرند.

استفاده از ترکیب‌های متفاوت فاکتورها به پاسخ‌های متفاوتی می‌انجامد بنابراین لازم است ترکیب بهینه‌ای از فاکتورها که به ازای آن‌ها الگوریتم‌های MLR، LR و MARS بهترین عملکرد را داشته باشند، انتخاب شود. در تحقیق پیش رو از الگوریتم ژنتیک ارائه شده توسط حسنلو و صمدزادگان (۲۰۱۱) برای بهینه‌سازی مدل استفاده شده است. این الگوریتم به صورت ضمنی وابستگی بین فاکتورها را نیز در نظر می‌گیرد و با توجه به آن بهترین زیرمجموعه از فاکتورها را انتخاب می‌کند. (Mahdipour & Dadkhah, ۲۰۱۴; Hong et al., ۲۰۱۸) در این تحقیق از RMSE نرمال شده که رابطه آن به صورت زیر است، به عنوان تابع برازش الگوریتم ژنتیک استفاده شده است:

$$NRMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y - \hat{y})^2} / \sigma_{\hat{y}}, \quad (14)$$

که در آن N تعداد داده‌ها، y مقدار واقعی متغیر وابسته، \hat{y} مقدار برآورد شده آن و $\sigma_{\hat{y}}$ انحراف معیار مقادیر برآورد شده متغیر وابسته است. شکل ۹ شبه کد پیشنهادی این الگوریتم برای تحقیق حاضر را نشان می‌دهد.

```

Inputs: Fire variable, Explanatory variables (factors), Parameters of the Genetic Algorithm.
Outputs: The factors with the most effects on the fire.
1. Generate random population of factors;
2. for  $G = 1$  to  $G_{max}$ 
3. Calculate the fitness of each individual;
4. if  $G = G_{max}$ 
5. return the factor-IDs; the fitness value; the mean and the best of the fitness values;
6. break;
7. end;
8. Elitism: Select the best individual of  $G$  and move it directly to  $G+1$ ;
9.  $S \leftarrow$  Select two individuals / parents of  $G$  at random, based on 'pair-wise tournament selection' to construct  $G+1$ ;
10. if  $rand \leq crossover\_rate$  //  $rand$  is a generated uniformly distributed random number value
11.  $S \leftarrow$  One-point crossover;
12. end;
13. for each individual in  $S$ 
14. if  $rand \leq mutation\_rate$ 
15. Mutation;
16. end;
17. end;
18. Add  $S$  to  $G+1$ ;
19. Go to step 8 to construct  $N$  individuals for  $G+1$ ;
20. end;

```

شکل ۹: شبه کد الگوریتم ژنتیک پیشنهادی برای حل مسئله انتخاب فاکتورهای بهینه

شرح و تفسیر نتایج

در تحقیق پیش رو وابستگی آتش‌سوزی‌های جنگلی در منطقه مورد مطالعه را به ۱۴ فاکتور مختلف که در جدول ۱ ارائه شده‌اند بررسی نمودیم. این ۱۴ فاکتور با توجه به مطالعات انجام شده و تحقیقات صورت گرفته در گذشته و محدودیت‌های موجود در دسترسی به داده‌ها انتخاب شده‌اند. مشاهدات ما نیز به صورت دودویی (۱ برای آتش و ۰ برای غیر آتش) وارد الگوریتم شدند.

جدول ۱: فاکتورهای مورد مطالعه در این تحقیق

شماره	فاکتور	شماره	فاکتور	شماره	فاکتور
۱	حداکثر دما (°C)	۶	حداکثر سرعت باد (m/s)	۱۱	جهت شیب
۲	حداقل دما (°C)	۷	جنس خاک	۱۲	مقدار شیب
۳	متوسط دما (°C)	۸	کاربری زمین	۱۳	ارتفاع (m)
۴	میزان بارندگی (mm)	۹	فاصله از جاده‌ها (m)	۱۴	فاصله از مناطق مسکونی (m)
۵	جهت باد غالب	۱۰	فاصله از رودخانه‌ها (m)		

ترتیب فاکتورها در تشکیل هر کروموزم از الگوریتم ژنتیک به همان ترتیبی است که در جدول ۱ ارائه شده است. جداول ۲ تا ۴ به ترتیب مقادیر پارامترهای الگوریتم‌های MARS, LR و ژنتیک مورد استفاده در این تحقیق را نشان می‌دهند.

جدول ۲: پارامترهای الگوریتم LR مورد استفاده

Parameter	Value
Maximum Iteration	۱۰۰
Threshold	۰/۰۰۰۰۰۱

جدول ۳: پارامترهای الگوریتم MARS مورد استفاده

Parameter	Value
Maximum Number of Basis Functions	۲۱
GCV Penalty per Knot (d)	۴
Piecewise-cubic Type of Modelling	۱
Fastness Level of Piecewise-cubic	۲
Maximum Degree of Self Interactions	۱
Maximum Degree of Interactions	۶
Threshold	۰/۰۰۰۱

جدول ۴: پارامترهای الگوریتم ژنتیک مورد استفاده

Parameter	Value
Population Size	۱۰
Number of Generations	۱۰۰
Crossover Rate	۰/۸
Migration Fraction	۰/۲
Migration Interval	۲۰
Scale (Gaussian Mutation)	۰/۵
Shrink (Gaussian Mutation)	۰/۷
Elite Count	۱

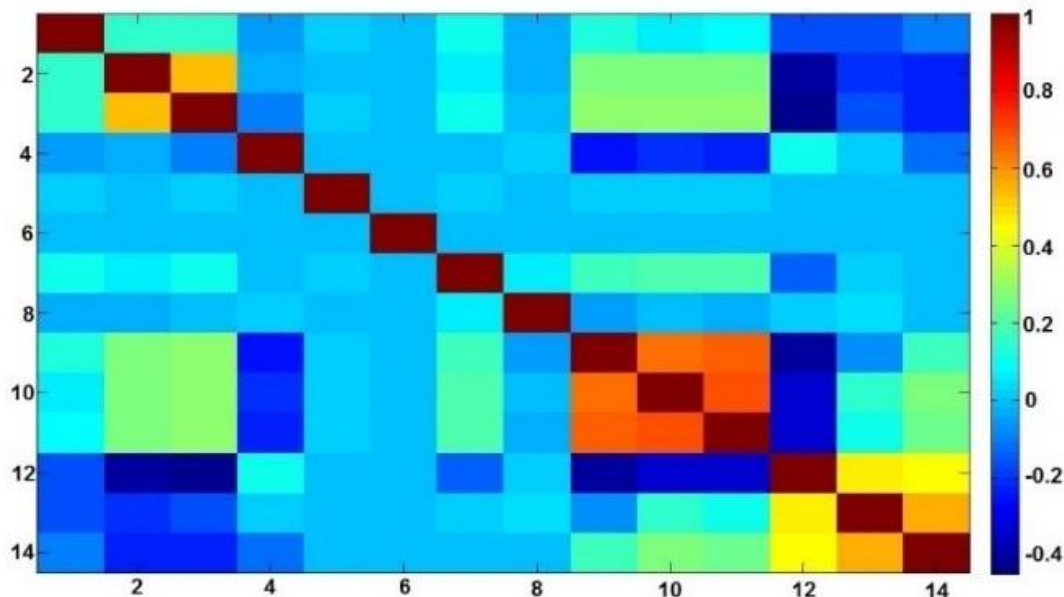
پیش از اجرای الگوریتم لازم است ابتدا همبستگی بین داده‌ها بررسی شود. بدین منظور از روابط زیر استفاده می‌شود (Dale, ۲۰۱۴):

$$Cov(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{n}, \quad (15)$$

$$r = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}, \quad (16)$$

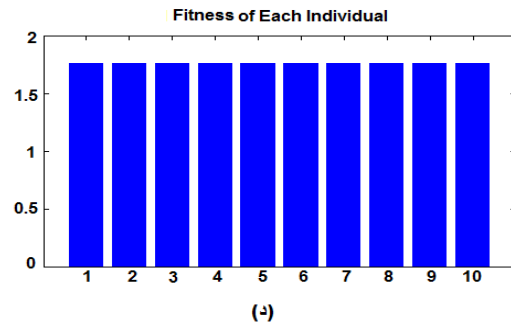
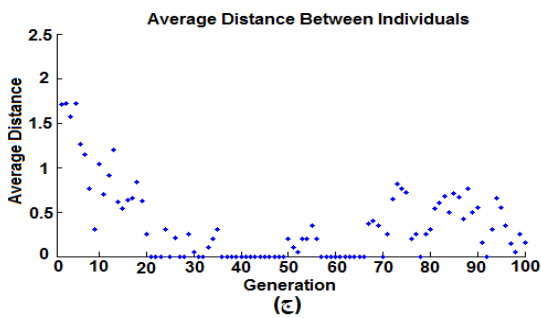
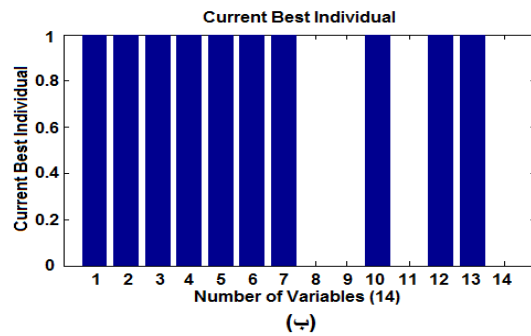
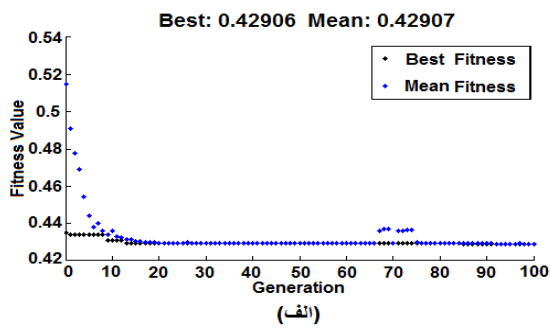
که در آن‌ها $Cov(X, Y)$ کوواریانس دو مجموعه داده X و Y ، \bar{X} و \bar{Y} میانگین این دو مجموعه داده، n تعداد داده‌های هر مجموعه و r ضریب همبستگی بین دو مجموعه داده با انحراف معیارهای σ_X و σ_Y است. تمام مقادیر ضریب همبستگی بین $-۰/۷$ و $۰/۷$ واقع است که نشان می‌دهد هیچ‌کدام از داده‌ها همبستگی قابل ملاحظه‌ای به سایر داده‌ها

ندارد. بنابراین از همه آن‌ها در الگوریتم‌ها استفاده می‌کنیم. شکل ۱۰ نمایشی از ماتریس همبستگی بین مجموعه داده‌های مورد استفاده در این مطالعه با ترتیبی مطابق جدول ۱ می‌باشد.

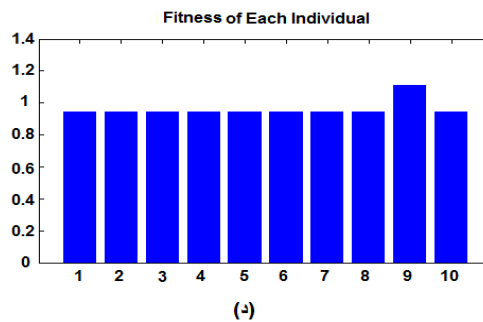
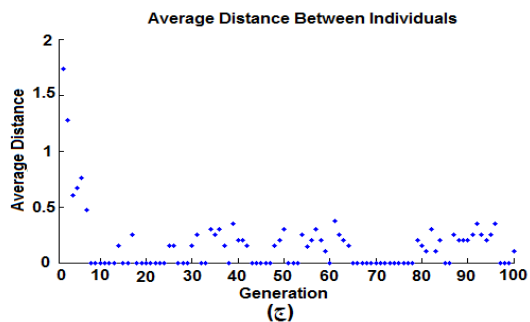
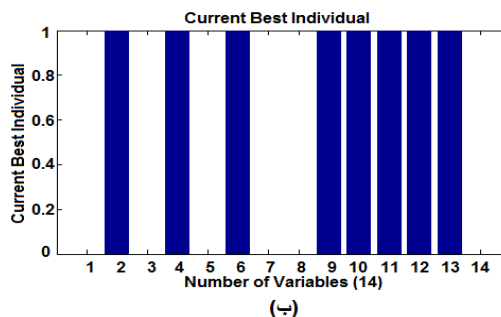
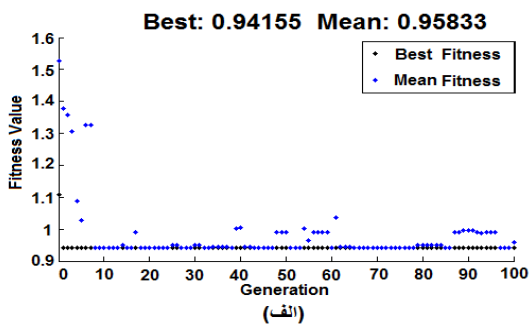


شکل ۱۰: ماتریس همبستگی بین داده‌ها

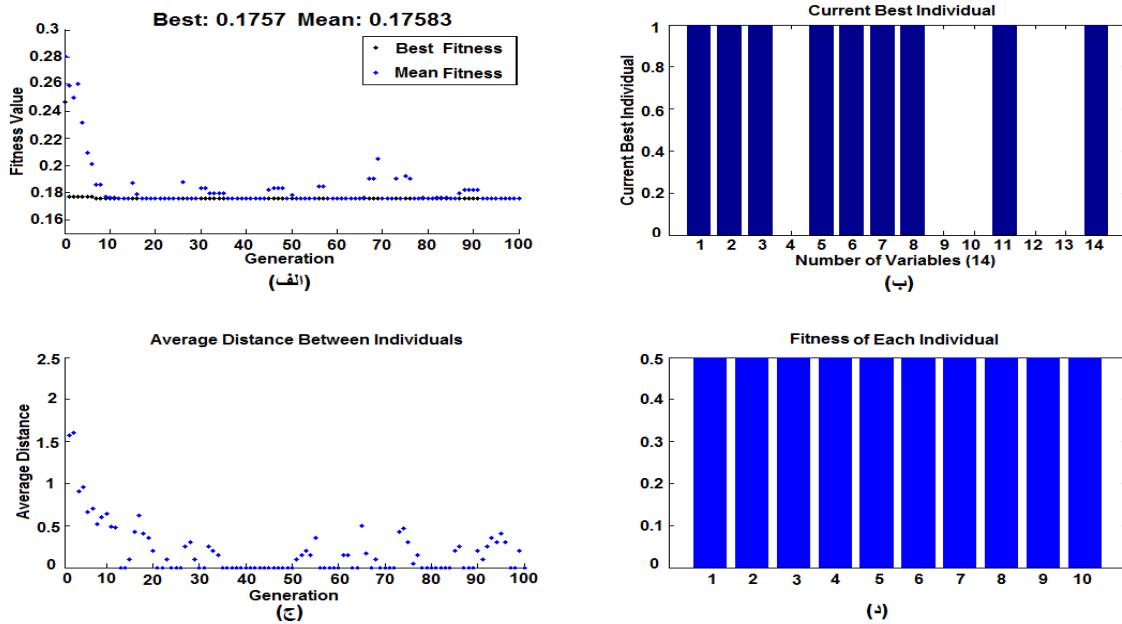
پس از اجرای الگوریتم ژنتیک با مدل‌های LR، MLR و MARS مقدار تابع برازش به ترتیب برابر با ۰/۴۲۹۱، ۰/۹۴۱۶ و ۰/۱۷۵۷ به دست آمد. در حالت اول ۱۰ فاکتور حداکثر دما، حداقل دما، متوسط دما، میزان بارندگی، جهت باد غالب، حداکثر سرعت باد، جنس خاک، فاصله از رودخانه‌ها، مقدار شیب و ارتفاع؛ در حالت دوم ۸ فاکتور حداقل دما، میزان بارندگی، حداکثر سرعت باد، فاصله از جاده‌ها، فاصله از رودخانه‌ها، جهت شیب، مقدار شیب و ارتفاع و در حالت سوم ۹ فاکتور حداکثر دما، حداقل دما، متوسط دما، جهت باد غالب، حداکثر سرعت باد، جنس خاک، کاربری زمین، جهت شیب و فاصله از مناطق مسکونی به عنوان فاکتورهای موثر شناخته شدند. میزان نوسان تابع برازش نیز در حالت استفاده از مدل MLR کمتر از سایر موارد است. شکل‌های ۱۱ تا ۱۳ نتایج حاصل از الگوریتم ژنتیک با پارامترهای مذکور در جدول ۴ را در هر سه حالت نمایش می‌دهند. ضرایب رگرسیون حاصل از روش‌های LR و MLR نیز در جدول ۵ ارائه شده است. لازم به ذکر است که ترتیب ضرایب در جدول ۵ با ترتیب فاکتورها در جدول ۱ یکی است و ضرایب مربوط به فاکتورهایی که موثر تشخیص داده نشدند با علامت (-) مشخص شده‌اند. شکل ۱۴ نیز نتایج روش MARS را نشان می‌دهد که در نهایت با ۸ تابع پایه و GCV برابر با ۰/۰۲۳۸ به پایان رسید.



شکل ۱۱: (الف) بهترین مقدار و میانگین مقادیر تابع برازش، (ب) بهترین افراد نسل آخر، (ج) میانگین فاصله بین افراد هر نسل و (د) مقدار برازش هر فرد در نسل آخر با استفاده از مدل MLR



شکل ۱۲: (الف) بهترین مقدار و میانگین مقادیر تابع برازش، (ب) بهترین افراد نسل آخر، (ج) میانگین فاصله بین افراد هر نسل و (د) مقدار برازش هر فرد در نسل آخر با استفاده از مدل LR



شکل ۱۳: (الف) بهترین مقدار و میانگین مقادیر تابع برازش، (ب) بهترین افراد نسل آخر، (ج) میانگین فاصله بین افراد هر نسل و (د) مقدار برازش هر فرد در نسل آخر با استفاده از مدل MARS

جدول ۵: ضرایب رگرسیون حاصل از روش‌های LR و MLR

LR	MLR	شماره
-	۰/۰۰۰۰۱	۱
-۰/۰۰۳۷	-۰/۰۰۰۰۷	۲
-	۰/۰۰۰۰۷	۳
۰/۰۰۲۰۷	۰/۰۰۲۷۲	۴
-	-۰/۰۱۹۹۱	۵
۰/۱۲۲۷۶	-۰/۰۰۰۰۷	۶
-	۰/۰۰۰۲۶	۷
-	-	۸
۰/۰۰۹۷۳	-	۹
-۰/۰۰۵۶۳	-۰/۳۸۴۵۱	۱۰
۸۰۰/۵۹۱۴	-	۱۱
۳۱/۳۹۸۱۶	۰/۶۹۵۰۷	۱۲
-۱۵۸۲/۱۲۱۸۱	۰/۰۱۲۱۶	۱۳
-	-	۱۴

علاوه بر RMSE یکی دیگر از معیارهای سنجش برازش ضریب تعیین R^2 (Rodriguez-Trejo et al., ۲۰۰۸) است که در جدول زیر مقدار آن را برای هر سه حالت مشاهده می‌کنید:

جدول ۶: ضریب تعیین R^2 مدل

MLR	LR	MARS	R^2
۰/۹۸۶۲	۰/۹۹۱۲	۰/۹۸۸۶	

همان‌طور که در شکل‌های ۱۱ تا ۱۳ نشان داده شده است، با معیار NRMSE روش MARS نسبت به دو روش دیگر عملکرد بهتری داشته است اما با توجه به جدول ۶ با معیار R^2 این روش LR است که به نتایج بهتری می‌انجامد؛ هر چند

^۱ Coefficient of determination

با این معیار اختلاف عملکرد سه روش بسیار کم است. در مجموع با توجه به نتایج به دست آمده می‌توان گفت که برای این مجموعه از داده‌ها روش MARS بهتر از دو روش دیگر عمل کرده است.

```

if (x7 <= 7.872) then f1_1 = 0
if (7.872 < x7 < 8.782) then begin
  p1_1 = (2*(8.782) + (7.872) - 3*(8.03)) / ((8.782) - (7.872))^2
  r1_1 = (2*(8.03) - (8.782) - (7.872)) / ((8.782) - (7.872))^3
  f1_1 = p1_1*(x7-(7.872))^2 + r1_1*(x7-(7.872))^3
end
if (x7 >= (8.782)) then f1_1 = x7 - (8.03)
if (x6 <= 3.708) then f1_2 = -(x6 - (3.891))
if (3.708 < x6 < 3.96) then begin
  p1_2 = (3*(3.891) - 2*(3.708) - (3.96)) / ((3.708) - (3.96))^2
  r1_2 = ((3.708) + (3.96) - 2*(3.891)) / ((3.708) - (3.96))^3
  f1_2 = p1_2*(x6-(3.96))^2 + r1_2*(x6-(3.96))^3
end
if (x6 >= (3.96)) then f1_2 = 0
BF1 = f1_1 * f1_2
if (x4 <= 11.63) then f2_3 = 0
if (11.63 < x4 < 12.04) then begin
  p2_3 = (2*(12.04) + (11.63) - 3*(11.97)) / ((12.04) - (11.63))^2
  r2_3 = (2*(11.97) - (12.04) - (11.63)) / ((12.04) - (11.63))^3
  f2_3 = p2_3*(x4-(11.63))^2 + r2_3*(x4-(11.63))^3
end
if (x4 >= (12.04)) then f2_3 = x4 - (11.97)
BF2 = BF1 * f2_3
if (x5 <= 186.1) then f3_1 = -(x5 - (207.9))
if (186.1 < x5 < 218.8) then begin
  p3_1 = (3*(207.9) - 2*(186.1) - (218.8)) / ((186.1) - (218.8))^2
  r3_1 = ((186.1) + (218.8) - 2*(207.9)) / ((186.1) - (218.8))^3
  f3_1 = p3_1*(x5-(218.8))^2 + r3_1*(x5-(218.8))^3
end
if (x5 >= (218.8)) then f3_1 = 0
if (x4 <= 12.04) then f3_2 = -(x4 - (12.11))
if (12.04 < x4 < 12.29) then begin
  p3_2 = (3*(12.11) - 2*(12.04) - (12.29)) / ((12.04) - (12.29))^2
  r3_2 = ((12.04) + (12.29) - 2*(12.11)) / ((12.04) - (12.29))^3
  f3_2 = p3_2*(x4-(12.29))^2 + r3_2*(x4-(12.29))^3
end
if (x4 >= (12.29)) then f3_2 = 0
if (x7 <= 7.185) then f3_3 = 0
if (7.185 < x7 < 7.872) then begin
  p3_3 = (2*(7.872) + (7.185) - 3*(7.714)) / ((7.872) - (7.185))^2
  r3_3 = (2*(7.714) - (7.872) - (7.185)) / ((7.872) - (7.185))^3
  f3_3 = p3_3*(x7-(7.185))^2 + r3_3*(x7-(7.185))^3
end
if (x7 >= (7.872)) then f3_3 = x7 - (7.714)
BF3 = f3_1 * f3_2 * f3_3
if (x3 <= 2) then f4_4 = 0
if (2 < x3 < 8.5) then begin
  p4_4 = (2*(8.5) + (2) - 3*(3)) / ((8.5) - (2))^2
  r4_4 = (2*(3) - (8.5) - (2)) / ((8.5) - (2))^3
  f4_4 = p4_4*(x3-(2))^2 + r4_4*(x3-(2))^3
end
if (x3 >= (8.5)) then f4_4 = x3 - (3)
BF4 = BF3 * f4_4
if (x3 <= 2) then f5_4 = -(x3 - (3))
if (2 < x3 < 8.5) then begin
  p5_4 = (3*(3) - 2*(2) - (8.5)) / ((2) - (8.5))^2
  r5_4 = ((2) + (8.5) - 2*(3)) / ((2) - (8.5))^3
  f5_4 = p5_4*(x3-(8.5))^2 + r5_4*(x3-(8.5))^3
end
if (x3 >= (8.5)) then f5_4 = 0
BF5 = BF3 * f5_4
if (x1 <= 1121) then f6_4 = 0
if (1121 < x1 < 3362) then begin
  p6_4 = (2*(3362) + (1121) - 3*(2242)) / ((3362) - (1121))^2
  r6_4 = (2*(2242) - (3362) - (1121)) / ((3362) - (1121))^3
  f6_4 = p6_4*(x1-(1121))^2 + r6_4*(x1-(1121))^3
end
if (x1 >= (3362)) then f6_4 = x1 - (2242)
BF6 = BF3 * f6_4
if (x1 <= 1121) then f7_4 = -(x1 - (2242))
if (1121 < x1 < 3362) then begin
  p7_4 = (3*(2242) - 2*(1121) - (3362)) / ((1121) - (3362))^2
  r7_4 = ((1121) + (3362) - 2*(2242)) / ((1121) - (3362))^3
  f7_4 = p7_4*(x1-(3362))^2 + r7_4*(x1-(3362))^3
end
if (x1 >= (3362)) then f7_4 = 0
BF7 = BF3 * f7_4
y = 0.01715 - 0.6491*BF1 - 38.87*BF2 + 3.228*BF3 - 0.4264*
BF4 - 1.824*BF5 - 0.002252*BF6 + 0.0003227*BF7

```

شکل ۱۴: توابع پایه حاصل از MARS و رابطه متغیر وابسته با آن‌ها

از بین فاکتورها تنها دو فاکتور حداقل دما و حداکثر سرعت باد در هر سه مدل به عنوان فاکتور موثر در آتش‌سوزی شناخته شدند؛ این در حالی است که فاکتورهای انسانی حاضر در این تحقیق یعنی کاربری زمین، فاصله از جاده‌ها و فاصله از مناطق مسکونی هر کدام تنها یک بار و سایر فاکتورها نیز دو بار به عنوان فاکتور موثر تشخیص داده شدند. آمارهای آتش‌سوزی‌های گذشته جنگل گلستان حاکی از این است که انسان نقش بسزایی در ایجاد حریق در این جنگل ایفا کرده است. بنابراین باید انتظار داشت که مدل‌های ما تا حدی با این آمارها نیز سازگاری داشته باشد. البته برای

دستیابی به نتایج بهتر باید تعداد فاکتورها به خصوص فاکتورهای انسانی را افزایش داد. حضور یا عدم حضور هر فاکتور در مسئله می‌تواند منجر به پاسخ‌های متفاوتی شود. بدین صورت که با افزودن فاکتورهای جدید به مسئله، علاوه بر این که خود این فاکتورها ممکن است به عنوان فاکتورهای مهم تشخیص داده شوند و میزان انطباق مدل بر داده‌ها و سایر خروجی‌ها را تغییر دهند، بر مهم شناخته شدن یا نشدن سایر فاکتورها نیز تأثیر می‌گذارند. بنابراین یکی از چالش‌های پیش رو جمع‌آوری داده‌های کافی و مناسب است. با این وجود حضور دو فاکتور انسانی کاربری زمین و فاصله از مناطق مسکونی در نتایج حاصل از مدل MARS نشان دهنده انطباق بیشتر آن به آمارها نسبت به دو روش دیگر است و می‌توان گفت نتایج آن با فرض اولیه این تحقیق همخوانی بیشتری دارد. در مجموع باید گفت که در مدیریت آتش‌سوزی جنگل گلستان باید به هر دو دسته فاکتورهای بیوفیزیکی و انسانی توجه کافی مبذول داشت.

نتیجه‌گیری

امروزه شناسایی فاکتورهای موثر بر آتش‌سوزی جنگل‌ها از اهمیت بسیار بالایی برخوردار است زیرا سالانه مساحت زیادی از جنگل‌های جهان بر اثر آتش‌سوزی نابود می‌شوند و تکرار این اتفاق در بلندمدت می‌تواند خسارات جبران‌ناپذیری بر زمین و ساکنین آن وارد کند. با شناسایی این فاکتورها می‌توانیم زمان‌ها و نقاط دارای ریسک بالای آتش‌سوزی را شناسایی نمائیم؛ به منظور پیش‌گیری از آتش‌سوزی با وضع قوانین و سیاست‌های مدیریتی کارآمد، آموزش به مردم و نظارت بیشتر در جهت مقابله با عوامل محرک آتش برآییم و در صورت بروز آتش‌سوزی تمهیدات لازم مانند ایجاد راه‌های دسترسی آسان به نقاط خطرناک و استقرار وسایل اطفاء حریق در نزدیکی این نقاط را بیندیشیم.

در این تحقیق سعی شد فاکتورهای موثر بر آتش‌سوزی‌های جنگل گلستان را شناسایی نمائیم و برای این منظور از سه روش رگرسیون خطی چندگانه، رگرسیون لجستیک و رگرسیون اسپلاین تطبیقی چندمتغیره در ترکیب با الگوریتم ژنتیک استفاده شد. نتایج این تحقیق نشان داد که هر دو دسته فاکتورهای بیوفیزیکی و انسانی در آتش‌سوزی‌های منطقه مورد مطالعه دارای تأثیر هستند. از این میان تنها فاکتورهای حداقل دما و حداکثر سرعت باد در هر سه حالت موثر شناخته شدند. روش رگرسیون اسپلاین تطبیقی چندمتغیره در مقایسه با دو روش دیگر عملکرد بهتری از خود نشان داد. مقدار RMSE نرمال شده این سه روش برابر ۰/۴۲۹۱ برای رگرسیون خطی چندگانه، ۰/۹۴۱۶ برای رگرسیون لجستیک و ۰/۱۷۵۷ برای رگرسیون اسپلاین تطبیقی چندمتغیره و مقدار R^2 آن‌ها نیز به ترتیب برابر ۰/۹۸۶۲، ۰/۹۹۱۲ و ۰/۹۸۸۶ به دست آمد.

در تحقیقات آینده پیشنهاد می‌شود تعداد فاکتورها افزایش یابد. در این زمینه می‌توان به فاکتورهای انسانی مانند تعداد گردشگر، سطح تحصیلی، فرهنگی و اقتصادی مردم، قوانین و سیاست‌های مدیریت جنگل و مدیریت بحران اشاره نمود. اگرچه جمع‌آوری این قبیل اطلاعات خود از چالش‌های پیش‌روست. از طرف دیگر می‌توان فاکتورهای بیوفیزیکی بیشتری نیز در نظر گرفت از جمله نوع و تراکم پوشش گیاهی و ویژگی‌های سوختی آن‌ها، رطوبت هوا، ساعات آفتابی و درصد بخار آب که متأسفانه به دلیل عدم دسترسی به داده‌های مناسب و کافی این امکان برای تحقیق حاضر فراهم نشد. همچنین می‌توان داده‌ها را در مقیاس زمانی بزرگتری تهیه نمود و خود زمان را به عنوان بعد چهارم وارد مسئله نمود. الگوریتم ارائه شده در این تحقیق می‌تواند در کنار روش‌های دیگر مانند شناسایی آتش به وسیله سنسور از دور برای دستیابی به نتایج بهتر به کار رود. همچنین می‌توان از آن به منظور پیش‌گیری یا شناسایی به موقع آتش در پیش‌بینی نحوه حرکت و گسترش آتش در جنگل بهره جست.

منابع

- Burgan, R., and R. Rothermel. ۱۹۸۴. *BEHAVE: Fire behavior prediction and fuel modelling system-fuel subsystem*. USDA Forest Service General Technical Report IN TI-۱۷. US Department of Agriculture, Forest Service, Intermountain Research Station, Ogden, Utah.
- Countryman, C.M. ۱۹۷۲. *Fire environment concept*. USDA Forest Service, Pacific Southwest Forest and Range Experiment Station, General Technical Report PSW-۷, Berkeley, CA.
- Craven, P., and G. Wahba. ۱۹۷۹. Smoothing noisy data with spline functions. Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik*, ۳۱: ۳۷۷-۴۰۳.
- Dale, P. ۲۰۱۴. *Mathematical Techniques in GIS, Second Edition*. CRC Press.
- Dickson, B.G.; J.W. Prather, Y. Xu, H.M. Hampton, E.N. Aumack, and T.D. Sisk. ۲۰۰۶. Mapping the probability of large fire occurrence in northern Arizona, USA. *Landscape Ecology*, ۲۱: ۷۴۷-۷۶۱.
- Forestry Canada Fire Danger Group Development and structure of the Canadian Forest Fire Behaviour Prediction System, Forestry Canada, Science and Sustainable Development Directorate. ۱۹۹۲. Report ST-X-۳. Ottawa, ON.
- Friedman, J.H. ۱۹۹۱. Multivariate Adaptive Regression Splines. *The Annals of Statistics*, ۱۹: ۱-۶۷.
- Hasanlou, M., and F. Samadzadegan. ۲۰۱۱. ICA/PCA base genetically band selection for classification of Hyperspectral images. *31st Asian Conference on Remote Sensing*.
- Hong, H.; P. Tsangaratos, I. Ilija, J. Liu, A.X. Zhu, and C. Xu. ۲۰۱۸. Applying genetic algorithms to set the optimal combination of forest fire related variables and model forest fire susceptibility based on data mining models. The case of Dayu County, China. *Science of the Total Environment*, ۶۳۰: ۱۰۴۴-۱۰۵۶.
- Kleinbaum, D.G., and M. Klein. ۱۹۹۴. *Logistic regression: A self learning text, Second Edition*. New York: Springer-Verlag Inc.
- Kong, S.G.; D. Jin, S. Li, and H. Kim. ۲۰۱۶. Fast fire flame detection in surveillance video using logistic regression and temporal smoothing. *Fire Safety Journal*, ۷۹: ۳۷-۴۳.
- Larjavaara, M.; J. Pennanen, and T. Tuomi. ۲۰۰۵. Lightning that ignites forest fires in Finland. *Agricultural and Forest Meteorology*, ۱۳۲: ۱۷۱-۱۸۰.
- Mahdipour, E., and C. Dadkhah. ۲۰۱۴. Automatic fire detection based on soft computing techniques: review from ۲۰۰۰ to ۲۰۱۰. *Artificial intelligence review*, ۴۲: ۸۹۵-۹۳۴.
- Martinez-Fernandez, J.; E. Chuvieco, and N. Koutsias. ۲۰۱۳. Modelling long-term fire occurrence factors in Spain by accounting for local variations with geographically weighted regression. *Natural Hazards and Earth System Science*, ۱۳: ۳۱۱-۳۲۷.
- Moritz, M.A.; J.E. Keeley, E.A. Johnson, and A.A. Schaffner. ۲۰۰۴. Testing a basic assumption of shrubland fire management: how important is fuel age? *Frontiers in Ecology and the Environment*, ۲: ۶۷-۷۲.
- Rodriguez-Trejo, D.A.; H.R. Maldonado, and J.H. Tchikoué. ۲۰۰۸. Factors affecting the accident rate of forest fires. *Ciencia Forestal en Mexico*, ۳۳: ۳۸-۵۷.
- Roman-Cuesta, R., and J. Martinez-Vilalta. ۲۰۰۶. Effectiveness of protected areas in mitigating fire within their boundaries: case study of Chiapas, Mexico. *Conservation Biology*, ۲۰: ۱۰۷۴-۱۰۸۶.
- Romero-Calcerrada, R.; C.J. Novillo, J.D.A. Millington, and I. Gomez-Jimenez. ۲۰۰۸. GIS analysis of spatial patterns of human-caused wildfire ignition risk in the SW of Madrid (Central Spain). *Landscape Ecology*, ۲۳: ۳۴۱-۳۵۴.
- Syphard, A.D.; V.C. Radeloff, N.S. Keuler, R.S. Taylor, T.J. Hawbaker, S.I. Stewart, and M.K. Clayton. ۲۰۰۸. Predicting spatial patterns of fire on a southern California landscape. *International Journal of Wildland Fire*, ۱۷: ۶۰۲-۶۱۳.
- West, A.M.; P.H. Evangelista, C.S. Jarnevich, and D. Schulte. ۲۰۱۸. A tale of two wildfires: testing detection and prediction of invasive species distributions using models fit with topographic and spectral indices. *Landscape Ecology*, ۳۳: ۹۶۹-۹۸۴.

Wotton, B., and D.L. Martell. ۲۰۰۵. A lightning fire occurrence model for Ontario. *Canadian Journal of Forest Research*, ۳۵: ۱۳۸۹-۱۴۰۱.

Yang, J.; H.S. He, S.R. Shifley, and E.J. Gustafson. ۲۰۰۷. Spatial patterns of modern period human-caused fire occurrence in the Missouri Ozark Highlands. *Forest science*, ۵۳: ۱-۱۵.

