

## تعیین فاکتورهای موثر بر آتش‌سوزی جنگل با استفاده از ترکیب رگرسیون وزن‌دار جغرافیایی و الگوریتم ژنتیک، مطالعه موردی: جنگل‌های استان گلستان

امین راعی<sup>۱</sup>، پرهام پهلوانی<sup>۲\*</sup>، مهدی حسنلو<sup>۲</sup>

۱- دانشجوی کارشناسی ارشد سیستم‌های اطلاعات مکانی، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی دانشگاه تهران

۲- استادیار دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی دانشگاه تهران

تاریخ دریافت مقاله: ۱۳۹۴/۰۶/۰۸ تاریخ پذیرش مقاله: ۱۳۹۵/۰۲/۱۵

### 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. با شناسایی این فاکتورها می‌توانیم نقاط و زمان‌های پرخطر برای آتش‌سوزی را شناسایی نمائیم و با وضع قوانین و سیاست‌های کارآمد در مدیریت جنگل و نظارت بیشتر در جهت مقابله با فاکتورهای محرک آتش برآییم. در این تحقیق، فاکتورهای موثر بر آتش‌سوزی‌های جنگل گلستان شناسایی شده‌اند. در این راستا، از روش رگرسیون وزن‌دار جغرافیایی (GWR) در ترکیب با الگوریتم ژنتیک استفاده شد. روش ترکیبی پیشنهادی روش مناسبی برای مسائل رگرسیون مکانی است زیرا فاکتورهای موثر با لحاظ نمودن دو ویژگی خاص داده‌های مکانی یعنی خودهمبستگی و ناپایداری مکانی شناسایی گشتند. در این تحقیق، از داده‌های سه آتش‌سوزی مختلف و دو روش وزن‌دهی گوسین و مکعبی سه‌گانه در GWR استفاده گشت. مقدار  $R^2$  با استفاده از هسته گوسین ۰/۹۵۳۸، ۰/۹۹۹۰ و ۰/۹۹۰۳ و با استفاده از هسته مکعبی سه‌گانه ۰/۹۹۳۱، ۰/۹۹۹۹ و ۰/۹۹۸۰ به دست آمد. نتایج تحقیقات نشان داد که هر دو دسته فاکتورهای بیوفیزیکی و انسانی تأثیر بسزایی در آتش‌سوزی‌های مناطق مورد مطالعه داشتند. از فاکتورهای بیوفیزیکی ارتفاع، جهت شیب، حداقل دما، متوسط دما و از فاکتورهای انسانی کاربری زمین و فاصله از مناطق مسکونی در بیشتر حالت‌ها موثر شناخته شدند. همچنین با استفاده از هسته مکعبی سه‌گانه برای وزن‌دهی در GWR، نتایج دقیق‌تر و مناسب‌تری به دست آمد.

**کلیدواژه‌ها:** آتش‌سوزی جنگل، رگرسیون وزن‌دار جغرافیایی، الگوریتم ژنتیک، جنگل گلستان.

\*نویسنده مسئول: دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی دانشگاه تهران.

تلفن: ۶۱۱۱۴۵۲۴

Email: pahlavani@ut.ac.ir

## ۱- مقدمه

امروزه، آتش‌سوزی جنگل‌ها در زمره مهم‌ترین و خطرناک‌ترین بلاهای طبیعی دنیا به شمار می‌آید و شناسایی فاکتورهای موثر بر ایجاد و گسترش آن به منظور تشخیص سریع آتش و واکنش بدان ضروری به نظر می‌رسد. تاکنون تحقیقات متعددی در این زمینه صورت گرفته است. در بیشتر موارد انسان عامل ایجاد آتش بوده است [۱، ۲ و ۳] اما عوامل انسانی موثر بر آتش‌سوزی جنگل‌ها هنوز به درستی شناسایی نشده است.

دو دسته اصلی فاکتورهای موثر بر آتش‌سوزی جنگل که در تحقیقات گذشته مورد بحث قرار گرفته شامل فاکتورهای بیولوژیکی- فیزیکی یا به‌طور خلاصه بیوفیزیکی و فاکتورهای اجتماعی- اقتصادی یا انسانی بوده‌اند. فاکتورهای بیوفیزیکی مانند نرخ بارگذاری سوخت تنها فاکتورهایی بودند که در مدیریت بحران در نظر گرفته می‌شدند [۴] در حالی که انسان با ایجاد یا خاموش کردن آتش تأثیر مستقیمی در روند آتش‌سوزی جنگل‌ها دارد، بنابراین باید تمامی این فاکتورها در فرآیند مدیریت آتش در نظر گرفته شوند [۵]؛ اگرچه فاکتورهای انسانی مانند کاربری زمین، فرآیندهای مدیریت سوخت، محدودیت‌های قانونی، میزان دسترسی به جنگل، عوامل فرهنگی، اجتماعی و اقتصادی گاه می‌توانند بسیار پیچیده باشند [۶ و ۷]. همچنین مطابق تحقیقات انجام شده توسط [۸]، فاکتورهای بیوفیزیکی به تنهایی نمی‌توانند الگوی آتش‌سوزی جنگل‌ها را تعیین کنند و باید با سایر فاکتورها ترکیب شوند.

اگرچه به طور کلی آتش‌سوزی رفتاری تصادفی دارد [۹] ولی [۱۰] آن را به صورت تابعی از سه متغیر در نظر گرفت، شامل (الف) شرایط جوی مانند دما، رطوبت، میزان بارندگی، جهت عمده وزش باد و سرعت حداکثر آن؛ (ب) ویژگی‌های سوختی مانند

نوع مواد سوختنی، نرخ بارگذاری سوخت، رطوبت مواد سوختنی و اشتعال‌پذیری آن‌ها؛ و (ج) توپوگرافی منطقه مانند ارتفاع، شیب و جهت شیب. بنابراین بیشتر تحقیقات صورت گرفته بر فاکتورهای بیوفیزیکی تمرکز نموده و کمتر به عوامل انسانی توجه کرده‌اند [۱۱، ۱۲ و ۱۳]. مدل‌های زیادی برای پیش‌بینی گسترش آتش در شرایط بیوفیزیکی متفاوت ارائه شده [۱۴ و ۱۵] و مدل‌های دیگری نیز برای شناسایی مناطق مستعد آتش‌سوزی به کار گرفته شده‌اند [۱۶ و ۱۷]. بسیاری از تحقیقات انجام شده در گذشته بر پایه نقطه شروع آتش‌سوزی بوده‌اند [۴، ۹ و ۱۸] و برخی دیگر از آن‌ها بر پایه سایر نقاط سوخته شده انجام شده‌اند [۱۹ و ۲۰] اما برای رسیدن به بهترین جواب باید تمامی این فاکتورها را به طور همزمان در نظر گرفت [۴ و ۲۱].

[۲۲] عنوان نمود که در مکزیک سرعت حداکثر باد، تعداد تولیدکنندگان و سطح تحصیلات آن‌ها بیشترین تأثیر را در آتش‌سوزی جنگل‌ها داشته است و [۲۳] نشان داد که در منطقه مورد مطالعه آن‌ها ارتباط معناداری بین آتش‌سوزی جنگل‌ها، تراکم جاده‌ها و کاربری کشاورزی وجود دارد. [۹] با استفاده از آنالیزهای آماری پارامتریک و غیر پارامتریک نشان داد که در منطقه مورد مطالعه آن‌ها احتمال وقوع آتش‌سوزی در محل‌های قابل دسترسی توسط انسان‌ها، در شیب‌های متوسط و ارتفاعات بالاتر و در جنگل‌های کاج و بلوط بیشتر است، که این نشان‌دهنده تأثیر توأمان فاکتورهای انسانی و بیوفیزیکی است. [۲۴] الگوی مکانی احتراق و تکرار آتش‌سوزی‌ها را با استفاده از مدل رگرسیون منطقی چندگانه<sup>۱</sup> برای فاکتورهای انسانی و

<sup>1</sup> Multiple logistic regression

تأثیر بیشتری هستند. ترکیب‌های متفاوت فاکتورها به پاسخ‌های متفاوتی می‌انجامد بنابراین یکی از چالش‌های پیش رو یافتن ترکیب بهینه فاکتورهاست که بدین منظور از الگوریتم ژنتیک استفاده نمودیم.

در ادامه تحقیق حاضر، در بخش ۲ به معرفی منطقه مورد مطالعه، داده‌های مورد استفاده، روش رگرسیون وزن‌دار جغرافیایی و الگوریتم ژنتیک مورد استفاده جهت یافتن بهترین ترکیب فاکتورها می‌پردازیم، در بخش ۳ نتایج به دست آمده ارائه و پیرامون آن‌ها بحث شده است و در نهایت، در بخش ۴ به نتیجه‌گیری از این تحقیق می‌پردازیم.

## ۲- مواد و روش‌ها

### ۲-۱- منطقه مورد مطالعه

جنگل گلستان از مهم‌ترین مناطق جنگلی ایران واقع در شرق استان گلستان و غرب استان خراسان شمالی می‌باشد. این جنگل از مناطق توریستی ایران می‌باشد و جاده‌های ارتباطی متعددی از میان آن عبور می‌کند. طبق آمار ثبت شده اکثر آتش‌سوزی‌های جنگل گلستان دارای منشأ انسانی بوده و در مجاورت این جاده‌ها رخ داده‌اند. منطقه مورد مطالعه این تحقیق (شکل ۱)، در محدوده  $37^{\circ}47' - 37^{\circ}20'$  شمالی و  $56^{\circ}12' - 55^{\circ}47'$  شرقی واقع شده است و مساحتی در حدود  $681/553$  کیلومتر مربع دارد. بخش عمده منطقه مورد مطالعه این تحقیق در استان خراسان شمالی واقع است و قسمت‌هایی از آن نیز در استان‌های گلستان و سمنان قرار دارد. این منطقه به گونه‌ای انتخاب شده است که به لحاظ اقلیمی دارای خصوصیات یکسانی باشد. شکل ۲ منطقه مورد مطالعه را در نقشه پهنه‌بندی اقلیمی ایران نشان می‌دهد. اطلاعات آتش این تحقیق مربوط به آتش‌سوزی‌های بزرگی است که در تاریخ‌های ۲۶ آبان ۱۳۸۹، ۸ فروردین ۱۳۹۰ و ۲۴ تیر ۱۳۹۰ در این منطقه اتفاق افتاده است.

رگرسیون پواسون یگانه و چندگانه<sup>۱</sup> برای فاکتورهای بیوفیزیکی در جنوب کالیفرنیا مدل‌سازی کرد و نتیجه گرفت که میزان احتراق بیشتر تحت تأثیر فاکتورهای انسانی و تکرار آتش‌سوزی‌ها تحت تأثیر فاکتورهای بیوفیزیکی است. [۲۵] با استفاده از مدل وزن‌های نشانگر<sup>۲</sup> (WofE) نشان داد که میزان دسترسی انسان‌ها به جنگل مهم‌ترین عامل در آتش‌سوزی‌های مرکز اسپانیا است. [۲۶] نشان داد که در جنگل‌های گالی‌پولی ترکیه نوع پوشش گیاهی، میزان شیب، جهت شیب و فاصله از جاده‌ها و مناطق مسکونی بیشترین تأثیر را در آتش‌سوزی جنگل‌ها دارد.

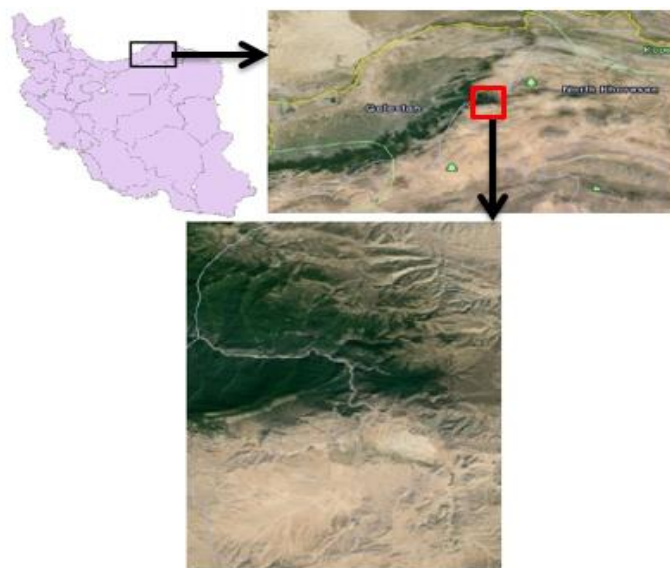
[۲۷] بیان کرد که فاکتورهای موثر بر آتش‌سوزی جنگل‌ها از نظر مکانی مستقل از هم نیستند و دارای خودهمبستگی مکانی می‌باشند و [۲۸] نشان داد که در این موارد رگرسیون وزن‌دار جغرافیایی<sup>۳</sup> (GWR) بهتر از سایر مدل‌های رگرسیون نتیجه می‌دهد. [۲۹] با استفاده از GWR نشان داد که کاربری زمین، نوع پوشش گیاهی و میزان بارندگی مهم‌ترین فاکتورهای موثر بر آتش‌سوزی جنگل‌ها در ایالت دورانگوی مکزیک بوده است و [۳۰] با استفاده از این مدل در منطقه مورد مطالعه خود فاکتورهای اقلیمی را موثرتر از فاکتورهای انسانی تشخیص داد.

هدف ما در این تحقیق تعیین فاکتورهای موثر بر آتش‌سوزی‌های بخشی از جنگل گلستان با استفاده از مدل رگرسیون وزن‌دار جغرافیایی بوده است. بیشتر آتش‌سوزی‌های اتفاق افتاده در این منطقه در سال‌های اخیر دارای منشأ انسانی بوده‌اند بنابراین هر دو فاکتور بیوفیزیکی و انسانی را در نظر گرفتیم و فرض ما این است که فاکتورهای انسانی دارای

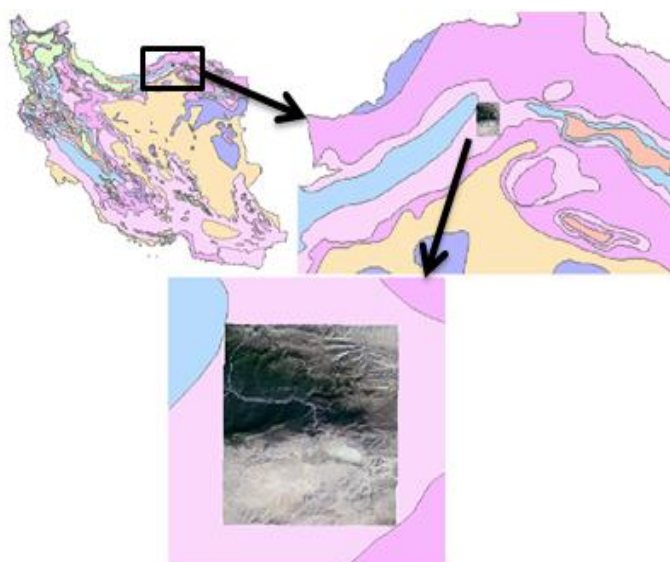
<sup>1</sup> Poisson univariate and multiple regressions

<sup>2</sup> Weights of Evidence

<sup>3</sup> Geographically Weighted Regression



شکل ۱: منطقه مورد مطالعه



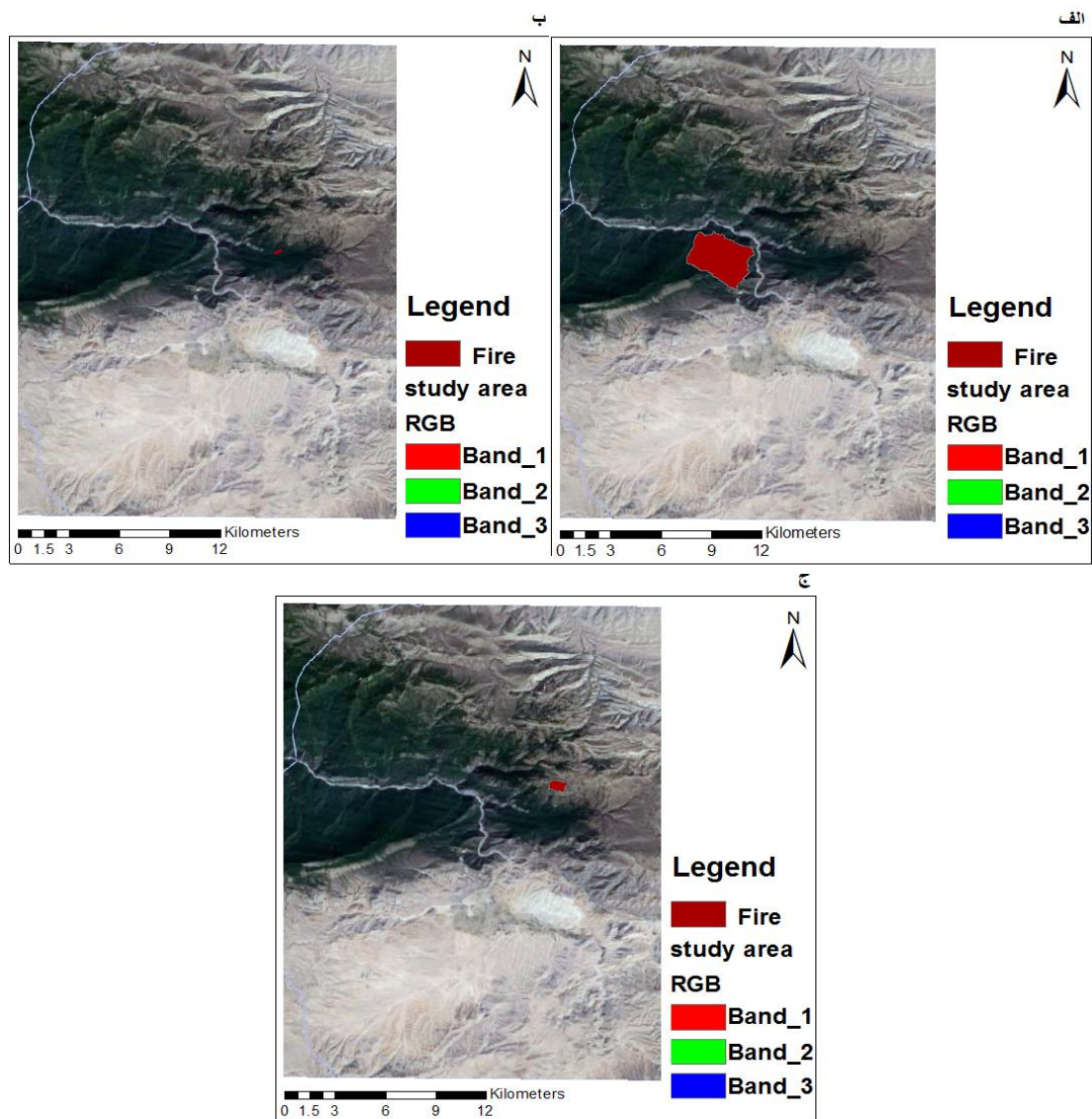
شکل ۲: منطقه مورد مطالعه در نقشه پهنه‌بندی اقلیمی ایران

## ۲-۲- داده‌ها

بخشی از داده‌های مورد نیاز این تحقیق شامل مناطق واقعی سوخته شده در آتش‌سوزی‌های ۲۶ آبان ۱۳۸۹، ۸ فروردین ۱۳۹۰ و ۲۴ تیر ۱۳۹۰ (شکل ۳) و داده‌های دیگری از قبیل مدل رقومی ارتفاعی<sup>۱</sup> (DEM) منطقه، شبکه راه‌ها، رودخانه‌ها، کاربری اراضی،

جنس خاک و مناطق مسکونی (شکل ۴) از سازمان نقشه‌برداری کشور و اداره کل منابع طبیعی استان‌های خراسان شمالی، گلستان و سمنان تهیه شده است. همچنین داده‌های ۵ ایستگاه هواشناسی سینوپتیک نزدیک به منطقه مورد مطالعه (شکل ۵) به همراه مختصات مکانی آن‌ها از طریق سازمان هواشناسی کشور به دست آمده است.

<sup>۱</sup> Digital Elevation Model

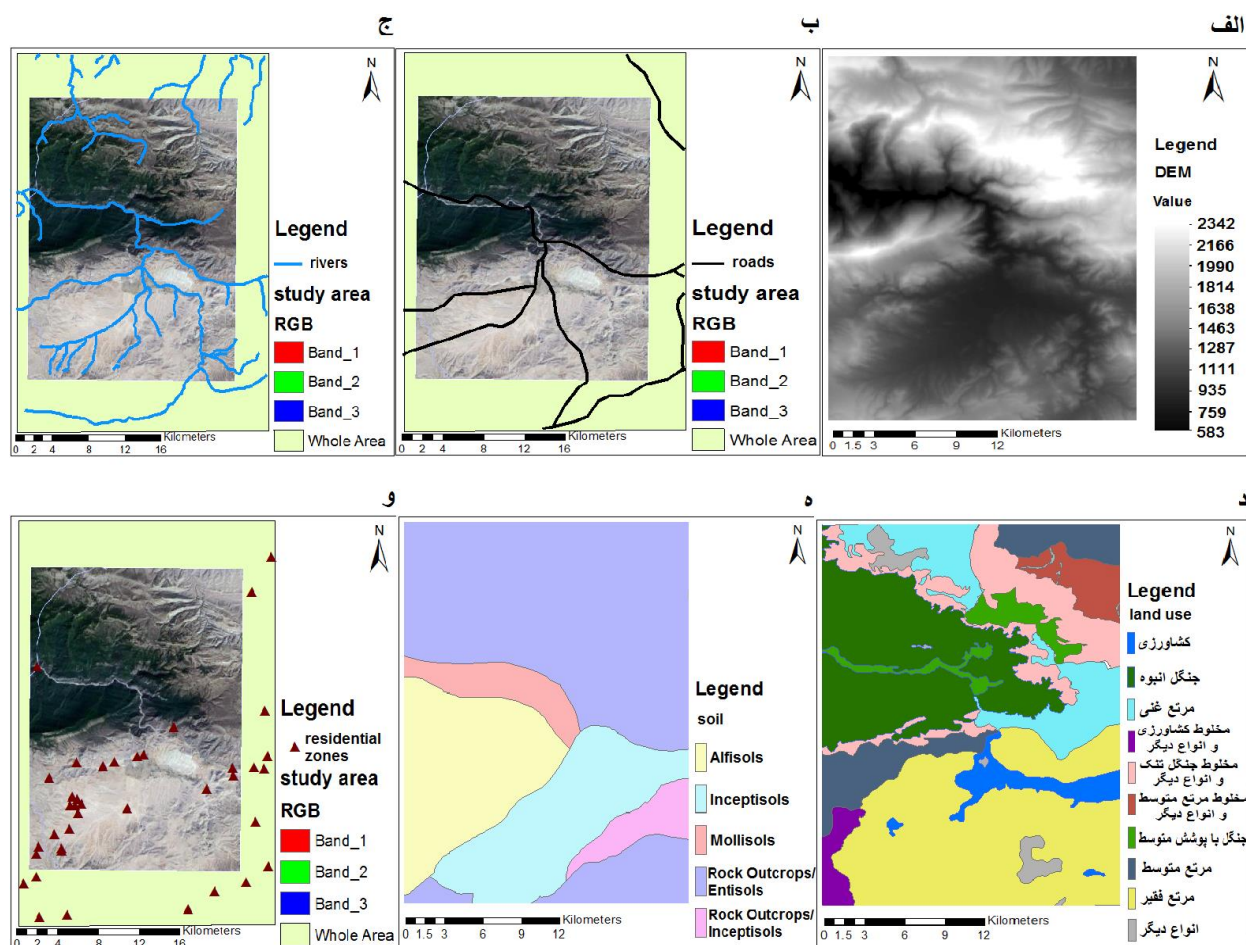


شکل ۳: آتش‌سوزی‌های اتفاق افتاده در منطقه مورد مطالعه در تاریخ‌های (الف) ۲۶ آبان ۱۳۸۹، (ب) ۸ فروردین ۱۳۹۰ و (ج) ۲۴ تیر ۱۳۹۰

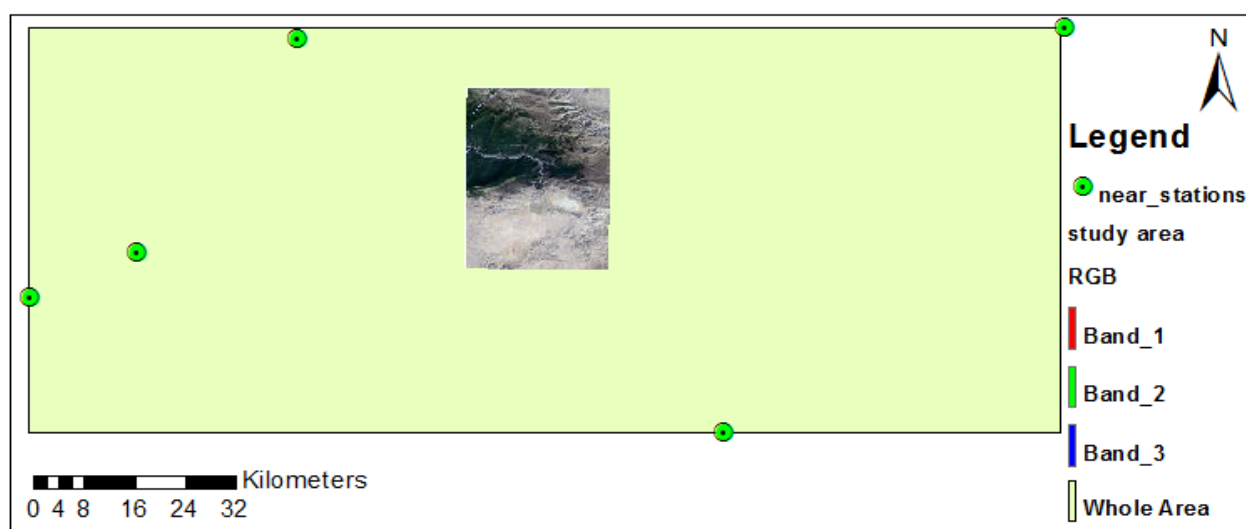
عمومی<sup>۱</sup> با مدل نمایی و قدرت تفکیک ۳۰ متر استفاده شد (شکل‌های ۶ تا ۸).

این داده‌ها شامل حداکثر، حداقل و متوسط دما، میزان بارش، سرعت حداکثر باد و جهت عمده وزش باد در آذر ماه سال ۸۹ می‌باشند. از آنجا که مقادیر این پارامترها در منطقه مورد 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

<sup>1</sup> Ordinary Kriging

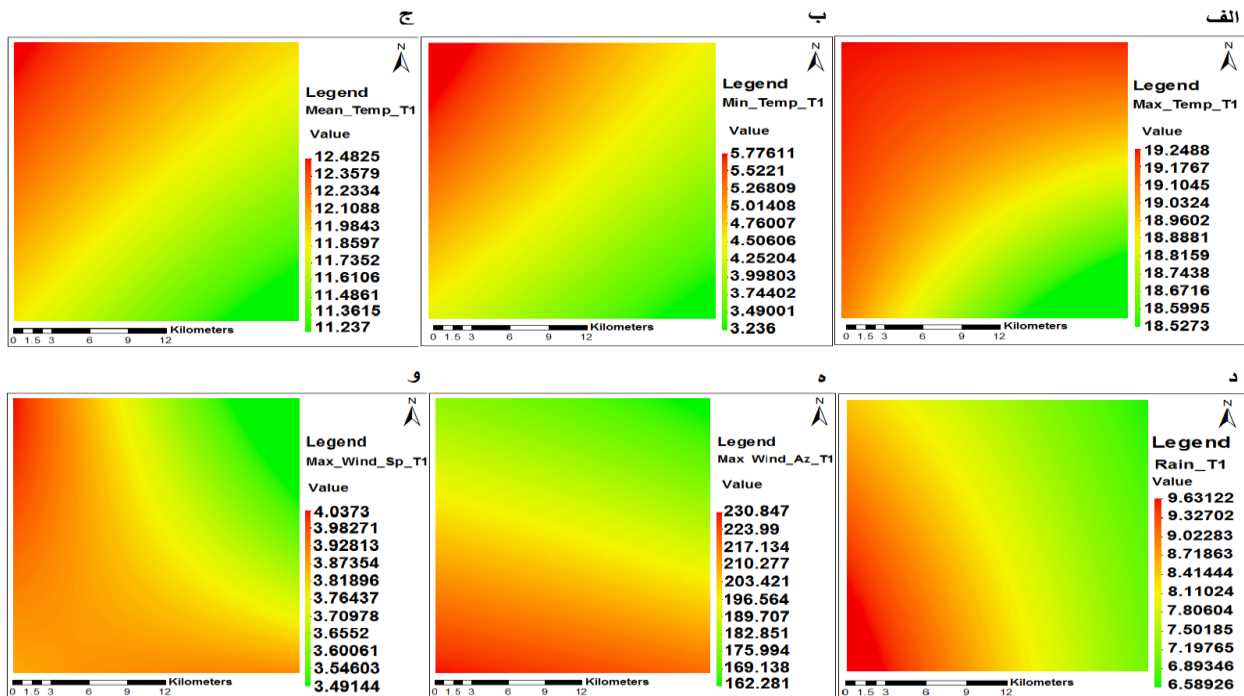


شکل ۴: لایه‌های اطلاعاتی مورد استفاده در این تحقیق (الف) مدل رقومی ارتفاعی منطقه، (ب) شبکه راه‌ها، (ج) رودخانه‌ها، (د) کاربری اراضی، (ه) جنس خاک و (و) مناطق مسکونی

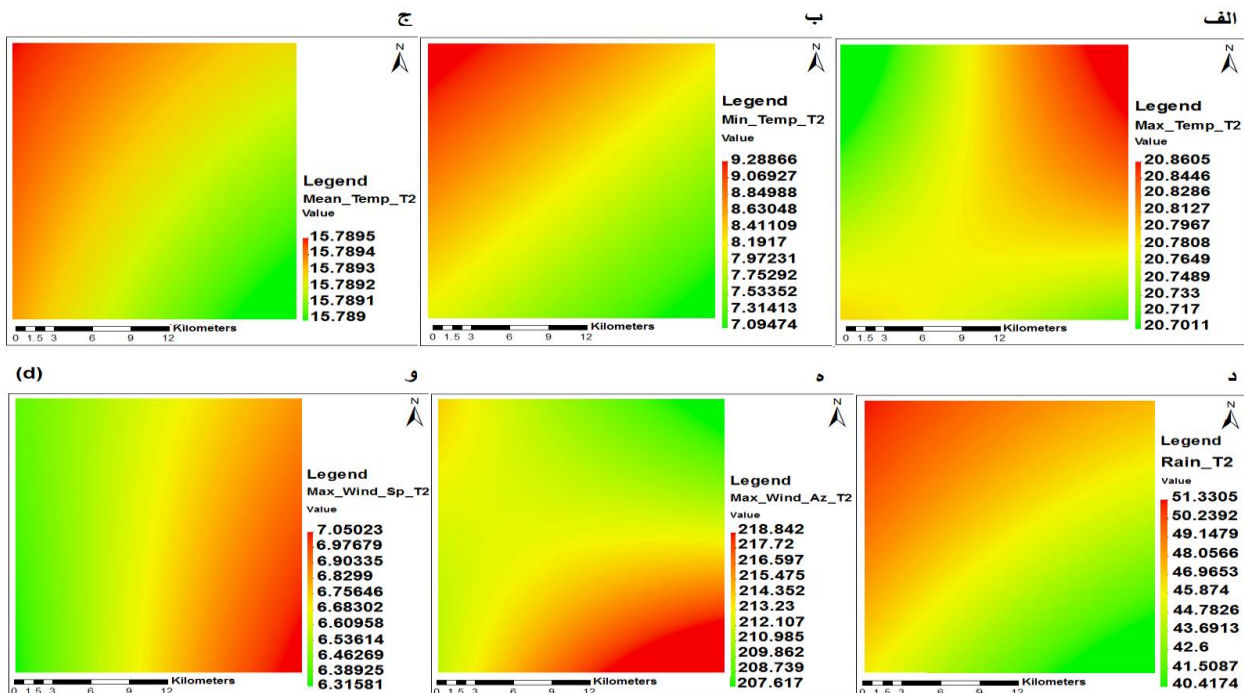


شکل ۵: ایستگاه‌های هواشناسی سینوپتیک نزدیک به منطقه مورد مطالعه

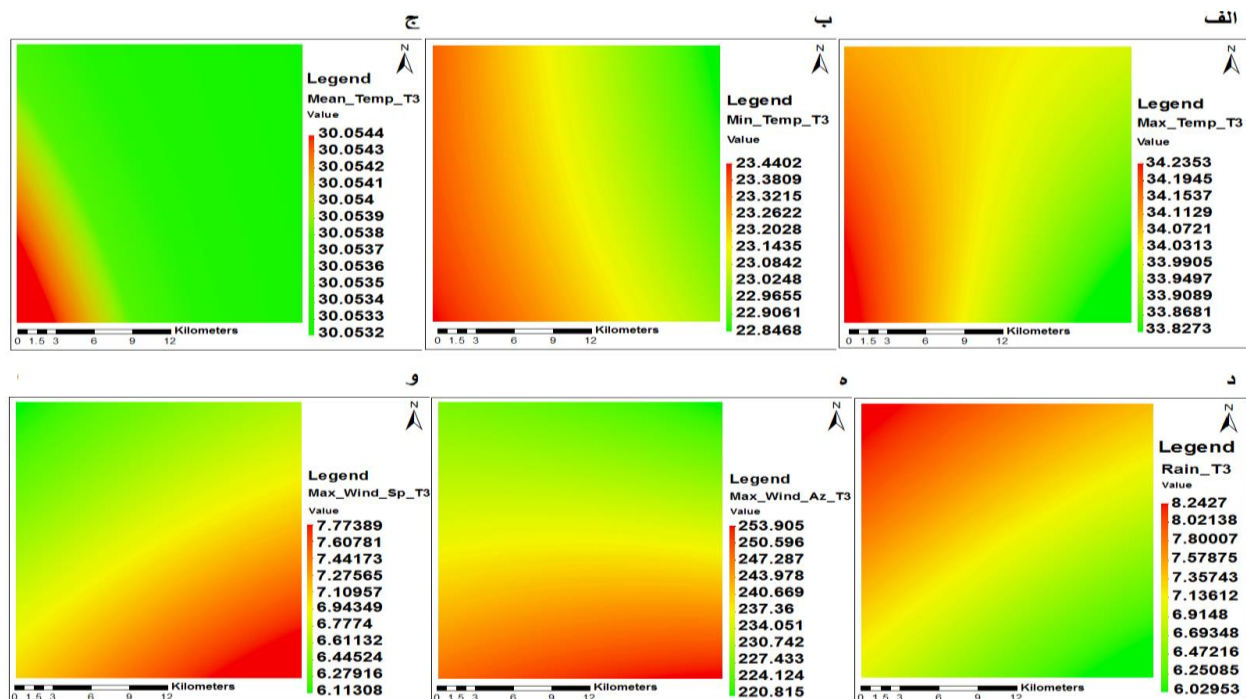




شکل ۶: رسترهای تولید شده از درون‌یابی (الف) حداکثر دما ( $^{\circ}\text{C}$ )، (ب) حداقل دما ( $^{\circ}\text{C}$ )، (ج) متوسط دما ( $^{\circ}\text{C}$ )، (د) میزان بارندگی (mm)، (ه) جهت باد غالب و (و) حداکثر سرعت باد (m/s) بین ایستگاه‌های هواشناسی نزدیک منطقه مورد مطالعه در تاریخ ۲۶ آبان ۱۳۸۹



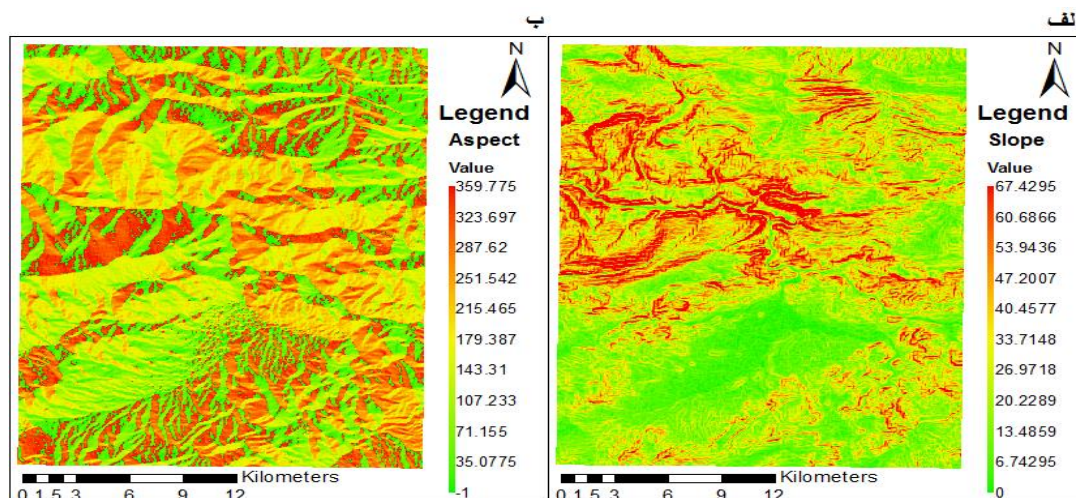
شکل ۷: رسترهای تولید شده از درون‌یابی (الف) حداکثر دما ( $^{\circ}\text{C}$ )، (ب) حداقل دما ( $^{\circ}\text{C}$ )، (ج) متوسط دما ( $^{\circ}\text{C}$ )، (د) میزان بارندگی (mm)، (ه) جهت باد غالب و (و) حداکثر سرعت باد (m/s) بین ایستگاه‌های هواشناسی نزدیک منطقه مورد مطالعه در تاریخ ۸ فروردین ۱۳۹۰



شکل ۸: رسترهای تولید شده از درون‌یابی (الف) حداکثر دما ( $^{\circ}\text{C}$ )، (ب) حداقل دما ( $^{\circ}\text{C}$ )، (ج) متوسط دما ( $^{\circ}\text{C}$ )، (د) میزان بارندگی (mm)، (ه) جهت باد غالب و (و) حداکثر سرعت باد (m/s) بین ایستگاه‌های هواشناسی نزدیک منطقه مورد مطالعه در تاریخ ۲۴ تیر ۱۳۹۰

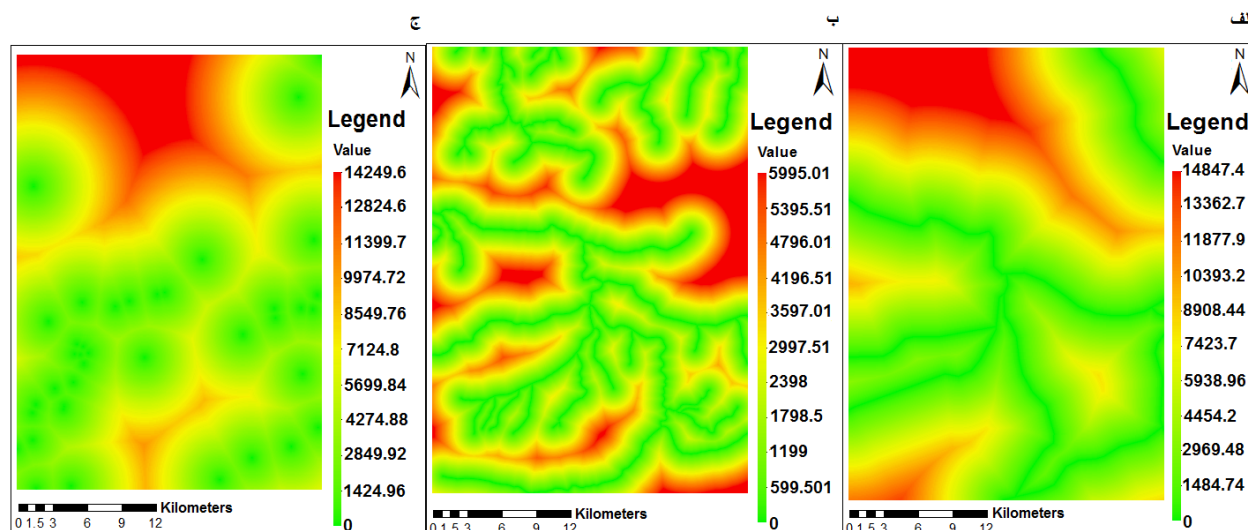
است (شکل ۹). لایه‌های اطلاعاتی راه‌ها، رودخانه‌ها و مناطق مسکونی دارای فرمت برداری بودند بنابراین با استفاده از آنالیز فاصله اقلیدسی، رسترهایی با قدرت تفکیک ۵ متر تولید گشت که هر سلول آن نشان‌دهنده فاصله از نزدیک‌ترین جاده، رودخانه یا منطقه مسکونی است (شکل ۱۰).

لایه‌های اطلاعاتی کاربری اراضی و جنس خاک 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 سازمان نقشه برداری 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



شکل ۹: لایه‌های اطلاعاتی (الف) شیب و (ب) جهت شیب



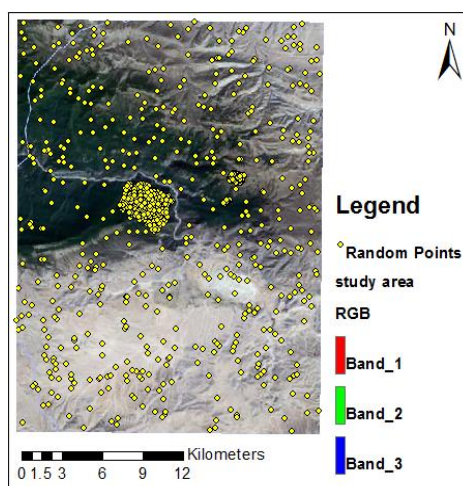


شکل ۱۰: لایه‌های اطلاعاتی فاصله از (الف) جاده‌ها (m)، (ب) رودخانه‌ها (m) و (ج) مناطق مسکونی (m)

## ۲-۳- رگرسیون وزن دار جغرافیایی

داده‌های مکانی دارای ویژگی‌های خاصی هستند. دو نمونه از این ویژگی‌ها عبارتند از (الف) خودهمبستگی مکانی<sup>۷</sup> بر اساس قانون تابلر که بیانگر رابطه معکوس وابستگی‌ها با فاصله است [۳۱] و (ب) ناپایداری مکانی<sup>۸</sup> که بیانگر تغییر خودهمبستگی مکانی در فضا و ناهمگونی محیط است. روش کمترین مربعات معمولی<sup>۹</sup> (OLS) نمی‌تواند با این دو ویژگی سازگار شود زیرا در این روش فرض می‌شود که داده‌ها کاملاً مستقل از هم هستند و محیط نیز همگون است بنابراین OLS بدون در نظر گرفتن وابستگی‌های مکانی برای تمامی نقاط منطقه یک دسته جواب ارائه می‌دهد. به همین دلیل در سال ۱۹۹۸، روش GWR توسط [۲۸] ارائه شد. در این روش، وابستگی‌های مکانی مشاهدات به صورت ماتریس‌های وزن در نظر گرفته می‌شود و به دلیل ناهمگونی محیط و وجود ناپایداری مکانی، ضرایب رگرسیون به صورت محلی و برای هر نقطه به طور جداگانه به دست می‌آید. رابطه کلی GWR بدین صورت است:

روش GWR با تعداد نقاط بالا بهتر نتیجه می‌دهد در حالی که تعداد نقاط مشاهداتی این تحقیق که همان ایستگاه‌های هواشناسی نزدیک به منطقه می‌باشند، بسیار کم است بنابراین ۱۰۰۰ نقطه به صورت تصادفی البته با تراکم بیشتر در محدوده آتش در منطقه مورد مطالعه با توزیع تصادفی یکنواخت تولید شد (شکل ۱۱). سپس، مقادیر همه لایه‌های اطلاعاتی موجود را در آن‌ها به دست آوردیم.



شکل ۱۱: نقاط تصادفی تولید شده در منطقه مورد مطالعه

<sup>۷</sup> Spatial autocorrelation

<sup>۸</sup> Spatial non-stationarity

<sup>۹</sup> Ordinary Least Squares

انتخاب پهنای باند مناسب است زیرا چنانچه این پارامتر بیش از حد بزرگ باشد، نتایج GWR به سمت OLS میل می‌کند و اگر بیش از حد کوچک انتخاب شود، واریانس نتایج به شدت افزایش می‌یابد [۳۳].

روش‌های مختلفی برای تعیین پهنای باند بهینه وجود دارد. یکی از آن‌ها روش اعتبارسنجی متقابل<sup>۱۲</sup> است که تابع آن بدین صورت است [۲۸]:

$$\sum_{i=1}^n [y_i - \hat{y}_i(h)]^2 \quad \text{رابطه (۶)}$$

که  $n$  تعداد مشاهدات،  $y_i$  مشاهده  $i$ ام و  $\hat{y}_i$  مقدار برآورد شده مشاهده  $i$ ام با استفاده از سایر مشاهدات است که خود تابعی از پارامتر پهنای باند است و هر پهنای باندی که این تابع را مینیمم کند به عنوان پهنای باند بهینه در نظر گرفته می‌شود.

خروجی GWR شامل پارامترهای متعددی است که از آن میان معمولاً پارامتر ضریب تعیین  $R^2$  برای سنجش خوبی برازش مدل به کار می‌رود [۳۳]. این پارامتر از رابطه (۷) بدست می‌آید [۳۴]:

$$R^2 = 1 - \frac{SS_E}{SS_T} \quad \text{رابطه (۷)}$$

$$SS_E = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad \text{رابطه (۸)}$$

$$SS_T = \sum_{i=1}^n (y_i - \bar{y})^2 \quad \text{رابطه (۹)}$$

که  $n$  تعداد مشاهدات،  $y_i$  مشاهده  $i$ ام،  $\hat{y}_i$  مقدار برآورد شده مشاهده  $i$ ام و  $\bar{y}$  میانگین مشاهدات است. در تحقیق پیش رو نیز از مقدار  $I-R^2$  به عنوان تابع برازش<sup>۱۳</sup> در الگوریتم ژنتیک استفاده شد که در بخش ۴-۲ به شرح آن می‌پردازیم.

همان‌طور که پیش‌تر ذکر شد، GWR توانایی مدل‌سازی تغییرات محلی و نالیستایی مکانی را دارد. آماره‌ای که می‌توان از آن برای تعیین

$$y = \beta_0(u, v) + \sum_{j=1}^p \beta_j(u, v)X_j + \varepsilon \quad \text{رابطه (۱)}$$

که در آن  $y$  متغیر وابسته،  $X_j$  متغیر مستقل  $j$ ام،  $p$  تعداد متغیرهای مستقل،  $\varepsilon$  باقیمانده مدل و  $\beta_j$  ضرایب رگرسیون هستند که خود تابعی از موقعیت نقاط مشاهداتی  $(u, v)$  می‌باشند. برخلاف OLS، GWR یک روش سرشکنی وزن‌دار است و ضرایب رگرسیون را از طریق رابطه زیر بدست می‌آورد:

رابطه (۲)

$$\hat{\beta}(u, v) = (X^T W(u, v) X)^{-1} X^T W(u, v) y$$

که در آن  $W$  ماتریس وزن مشاهدات است که خود تابعی از موقعیت نقاط می‌باشد. این ماتریس یک ماتریس قطری شامل وزن‌های جغرافیایی به شکل رابطه (۳) است:

$$\begin{bmatrix} W_1(u, v) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & W_n(u, v) \end{bmatrix} \quad \text{رابطه (۳)}$$

تعیین وزن‌های جغرافیایی در GWR از اهمیت بالایی برخوردار است. تاکنون هسته‌های متعددی بدین منظور ارائه شده است. دو نمونه از این هسته‌ها که کارایی بالایی را ثابت کرده‌اند هسته گوسین و مکعبی سه گانه<sup>۱۴</sup> می‌باشند که به ترتیب در روابط (۴) و (۵) نشان داده شده‌اند [۳۲]:

$$W_{ij} = \varphi\left(\frac{d_{ij}}{\sigma h}\right) \quad \text{رابطه (۴)}$$

$$W_{ij} = \begin{cases} \left(1 - \left(\frac{d_{ij}}{h}\right)^3\right)^3, & d_{ij} \leq h \\ 0, & d_{ij} > h \end{cases} \quad \text{رابطه (۵)}$$

که  $W_{ij}$  وزن جغرافیایی مربوط به مشاهده  $j$ ام در نقطه  $i$ ام،  $\varphi$  تابع توزیع نرمال استاندارد،  $d_{ij}$  فاصله اقلیدسی دو نقطه  $i$  و  $j$ ،  $\sigma$  انحراف استاندارد مقادیر  $d_{ij}$  به ازای هر نقطه و  $h$  پارامتر پهنای باند می‌باشد. مسئله مهم در تعیین وزن‌های جغرافیایی

<sup>۱۲</sup> Cross Validation

<sup>۱۳</sup> Fitness function

<sup>۱۰</sup> Kernel

<sup>۱۱</sup> Tricube

تقریباً تصادفی باشند و مقدار این آماره نزدیک به صفر شود که در این صورت نشان‌دهنده توانایی بالای GWR در مدل‌سازی خودهمبستگی مکانی مشاهدات می‌باشد [۳۵].

## ۲-۴- انتخاب فاکتورهای بهینه

در این مرحله، نسبت به انتخاب زیرمجموعه‌ای از فاکتورها که به ازای آن‌ها الگوریتم GWR بهترین عملکرد را داشته باشد، اقدام می‌شود. در این تحقیق، از الگوریتم ژنتیک ارائه شده در [۳۶] برای انتخاب فاکتورهای بهینه 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. الگوریتم ژنتیک یک الگوریتم اکتشاف جستجوگرایانه<sup>۱۵</sup> بوده که بر اساس انتخاب طبیعی کار می‌کند. این اکتشاف اغلب برای حل مسائل جستجو و بهینه‌سازی به کار برده می‌شود. در الگوریتم ژنتیک اغلب یک جمعیت از راه‌حل‌های نماینده (کروموزوم) به سمت بهینه کلی حرکت می‌کنند. روند بهینه‌سازی از یک جمعیت (نسل) تشکیل شده از کروموزوم‌های تولید شده به صورت اتفاقی شروع می‌شود. در هر نسل تابع برازش برای تمامی کروموزوم‌های آن نسل محاسبه می‌شود. این تابع برازش معمولاً تابع هدف مسئله بهینه‌سازی است. همان‌طور که در بخش قبل بیان شد، تابع برازش الگوریتم ژنتیک پیشنهادی این تحقیق،  $I-R^2$  و هدف ما از بهینه‌سازی مینیمم نمودن آن است.

بر اساس مکانیسم نخبه‌گرایی<sup>۱۶</sup>، بهترین کروموزوم هر نسل به صورت مستقیم به نسل بعد منتقل می‌گردد. همچنین، کروموزوم‌های دارای مقدار تابع برازش کمتر به صورت اتفاقی در هر نسل

میزان تغییرپذیری ضرایب رگرسیون در منطقه مورد مطالعه استفاده نمود، انحراف استاندارد ضرایب رگرسیون است که با رابطه زیر بیان می‌شود [۲۸]:

$$v_j = \sqrt{\sum_i (\beta_{ij} - \beta_j)^2 / n} \quad \text{رابطه (۱۰)}$$

که  $\beta_{ij}$  ضریب رگرسیون فاکتور  $j$ ام در مشاهده  $i$ ام،  $\beta_j$  میانگین ضریب رگرسیون فاکتور  $j$ ام در کل مشاهدات و  $n$  تعداد مشاهدات می‌باشد.

همچنین مقادیر خطای جذر میانگین مربعات<sup>۱</sup> (RMSE) و RMSE نرمال شده (NRMSE) باقیمانده‌های مدل که از روابط زیر به دست می‌آیند، به منظور سنجش توزیع این باقیمانده‌ها به کار می‌روند:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2} \quad \text{رابطه (۱۱)}$$

$$NRMSE = RMSE / \sigma_{\hat{y}} \quad \text{رابطه (۱۲)}$$

که در آن‌ها  $n$  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.

آماره دیگری که غالباً به منظور ارزیابی GWR مورد استفاده قرار می‌گیرد، شاخص موران<sup>۲</sup> است که رابطه آن به صورت زیر می‌باشد [۳۴]:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} z_i z_j}{S_0 \sum_{i=1}^n z_i^2} \quad \text{رابطه (۱۳)}$$

که  $n$  تعداد مشاهدات،  $w_{ij}$  وزن بین مشاهدات  $i$  و  $j$ ،  $z_i$  و  $z_j$  به ترتیب انحراف مشاهدات  $i$  و  $j$  از میانگین مشاهدات و  $S_0$  مجموع کلیه وزن‌هاست. از این آماره برای تعیین خودهمبستگی مکانی باقیمانده‌های مدل استفاده می‌شود. چنانچه فرآیند سرشکنی خطاها به درستی انجام شود می‌بایست باقیمانده‌های مدل دارای توزیع مکانی

<sup>15</sup> Search Heuristic

<sup>16</sup> Elitism

<sup>1</sup> Root Mean Squares Error

<sup>15</sup> Moran's I

در نظر گرفته شده‌اند. 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) علاوه بر مختصات نقاط که برای محاسبه وزن‌های جغرافیایی لازم می‌باشند به عنوان دیگر ورودی‌های الگوریتم برای حل مساله مد نظر قرار گرفتند.

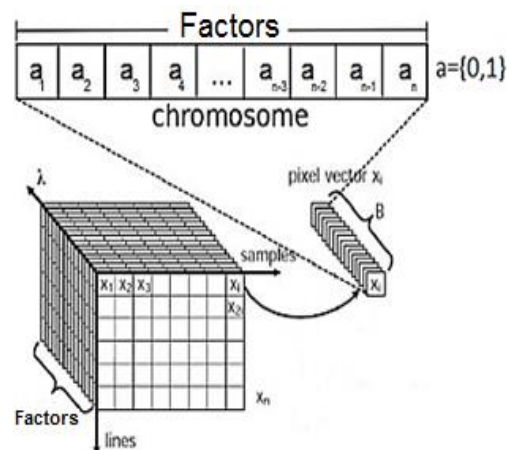
پیش از اجرای الگوریتم لازم است ابتدا همبستگی بین داده‌ها بررسی شود. بدین منظور از روابط زیر استفاده می‌شود [۳۷]:

$$Cov(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{n} \quad \text{رابطه (۱۴)}$$

$$r = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} \quad \text{رابطه (۱۵)}$$

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  $\sigma_X$  و  $\sigma_Y$  است. تمام مقادیر ضریب همبستگی بین  $-0.7$  و  $0.7$  واقع است که نشان می‌دهد هیچ یک از داده‌ها همبستگی قابل ملاحظه‌ای به سایر داده‌ها ندارد. بنابراین از همه آن‌ها در الگوریتم GWR استفاده می‌کنیم. شکل‌های ۱۳ تا ۱۵ نمایشی از ماتریس همبستگی بین مجموعه داده‌های مورد استفاده در این مطالعه با ترتیبی مطابق جدول ۱ می‌باشد.

بر اساس مکانیسم انتخاب؛ انتخاب شده و ژنوم<sup>۱</sup> این کروموزوم‌ها با اعمال تغییراتی از سوی عملگرهای ژنتیک (ترکیب ضربی<sup>۲</sup> و جهش<sup>۳</sup>) به نسل بعد انتقال می‌یابند. این نسل جدید از کروموزوم‌های تولید شده در حلقه تکرار بعدی مورد استفاده قرار می‌گیرند. بدین ترتیب، زمانی الگوریتم ژنتیک خاتمه می‌یابد که یا به تعداد پیشینه نسل‌ها و یا به مقدار تابع برازش رضایت بخش رسیده باشد. بر این اساس، شکل ۱۲ نحوه کدگذاری الگوریتم ژنتیک پیشنهادی این تحقیق برای حل مسئله یافتن فاکتورهای بهینه را نشان می‌دهد به گونه‌ای که  $a_i$  به  $i$ -امین فاکتور اشاره دارد.



شکل ۱۲: نحوه کدگذاری هر کروموزوم الگوریتم ژنتیک پیشنهادی این تحقیق برای حل مسئله یافتن فاکتورهای بهینه

### ۳- پیاده‌سازی و ارزیابی

در این تحقیق وابستگی آتش‌سوزی‌های جنگلی در منطقه مورد مطالعه را به ۱۴ فاکتور مختلف که در جدول ۱ ارائه شده‌اند بررسی نمودیم. همان‌طور که پیش‌تر ذکر شد برای دستیابی به نتایج بهتر هر دو دسته فاکتورهای انسانی و بیوفیزیکی

<sup>17</sup> Genome

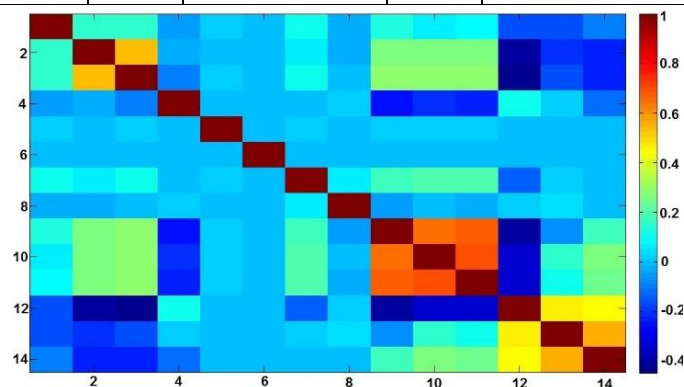
<sup>2</sup> Crossover

<sup>3</sup> Mutation

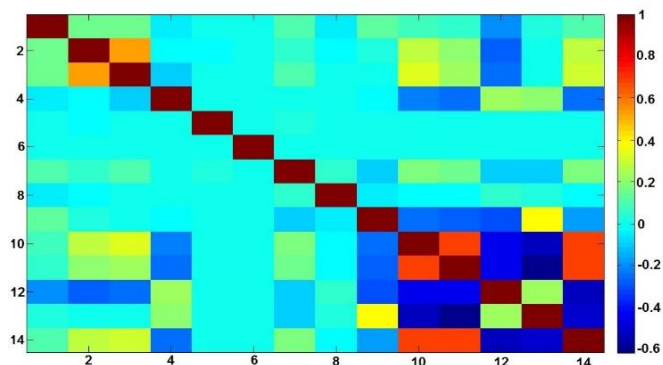


جدول ۱: فاکتورهای مورد مطالعه در این تحقیق

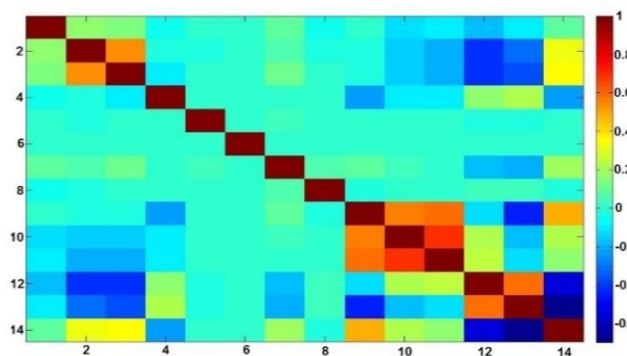
شماره	فاکتور	شماره	فاکتور	شماره	فاکتور
۱	فاصله از رودخانه‌ها (m)	۶	ارتفاع (m)	۱۱	متوسط دما ( $^{\circ}\text{C}$ )
۲	فاصله از جاده‌ها (m)	۷	مقدار شیب	۱۲	جهت باد غالب
۳	فاصله از مناطق مسکونی (m)	۸	جهت شیب	۱۳	حداکثر سرعت باد (m/s)
۴	جنس خاک	۹	حداکثر دما ( $^{\circ}\text{C}$ )	۱۴	میزان بارندگی (mm)
۵	کاربری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



شکل ۱۳: ماتریس همبستگی داده‌های ۲۶ آبان ۱۳۸۹



شکل ۱۴: ماتریس همبستگی داده‌های ۸ فروردین ۱۳۹۰



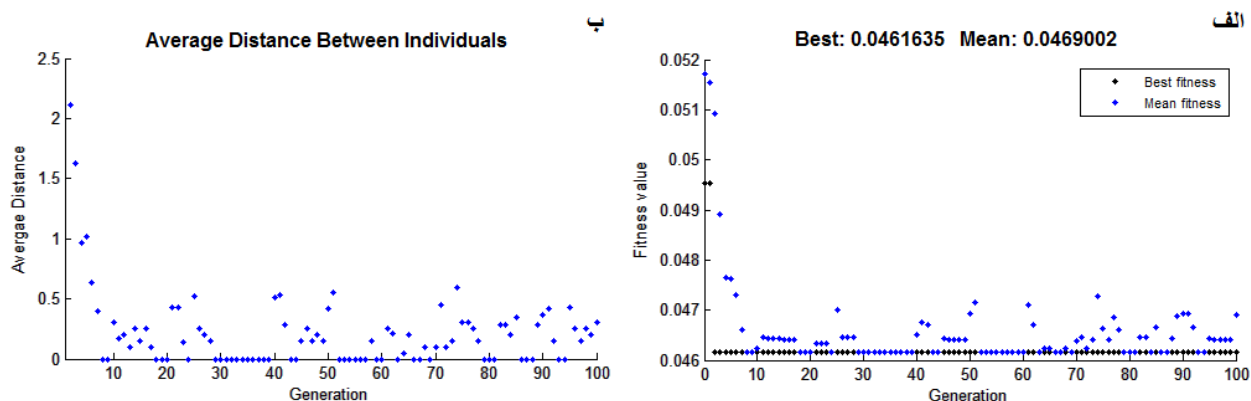
شکل ۱۵: ماتریس همبستگی داده‌های ۲۴ تیر ۱۳۹۰

مناطق مسکونی، کاربری زمین، مقدار شیب، جهت شیب، متوسط دما، حداکثر سرعت باد و با استفاده از هسته مکعبی سه‌گانه، ۶ فاکتور فاصله از جاده‌ها، جنس خاک، کاربری زمین، ارتفاع، متوسط دما، میزان بارندگی و سرانجام در آتش‌سوزی ۲۴ تیر ۱۳۹۰ با استفاده از هسته گوسین، ۷ فاکتور فاصله از مناطق مسکونی، کاربری زمین، ارتفاع، حداکثر دما، حداقل دما، متوسط دما، جهت باد غالب و با استفاده از هسته مکعبی سه‌گانه، ۷ فاکتور فاصله از مناطق مسکونی، جنس خاک، ارتفاع، مقدار شیب، جهت شیب، حداقل دما و حداکثر سرعت باد به عنوان فاکتورهای موثر شناخته شدند. این نتایج تا حد زیادی با ماتریس‌های همبستگی شکل‌های ۱۳ تا ۱۵ همخوانی دارد و در اکثر موارد فاکتورهای با همبستگی بالای ۰/۶ به صورت توأمان در خروجی‌ها ظاهر نشدند. شکل‌های ۱۶ تا ۲۱ نتایج حاصل از الگوریتم ژنتیک با پارامترهای مذکور در جدول ۲ را نمایش می‌دهند.

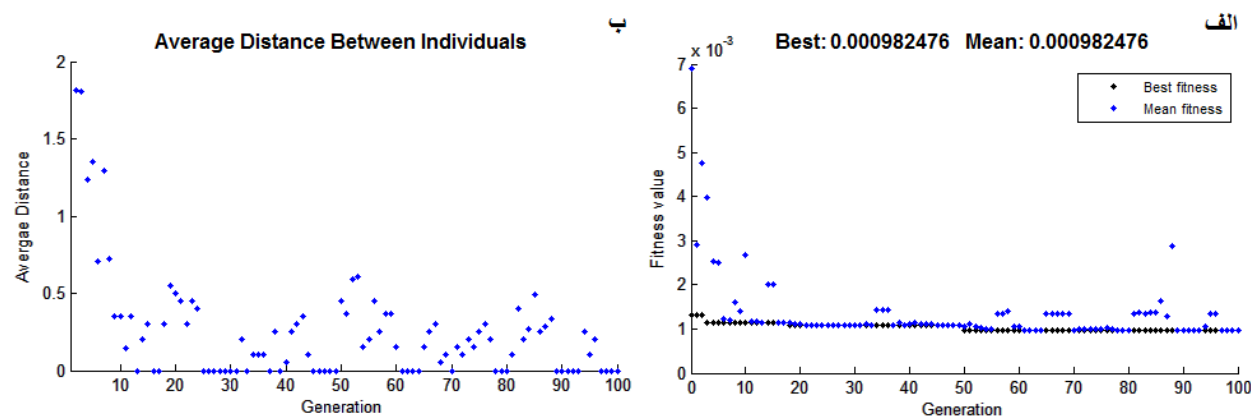
جدول ۲: پارامترهای الگوریتم ژنتیک مورد استفاده

Parameter	Value
Population size	10
Number of Generations	100
Crossover rate	0.8
Migration Fraction	0.2
Migration Interval	20
Scale (Gaussian Mutation)	0.5
Shrink (Gaussian Mutation)	0.7
Elite count	1

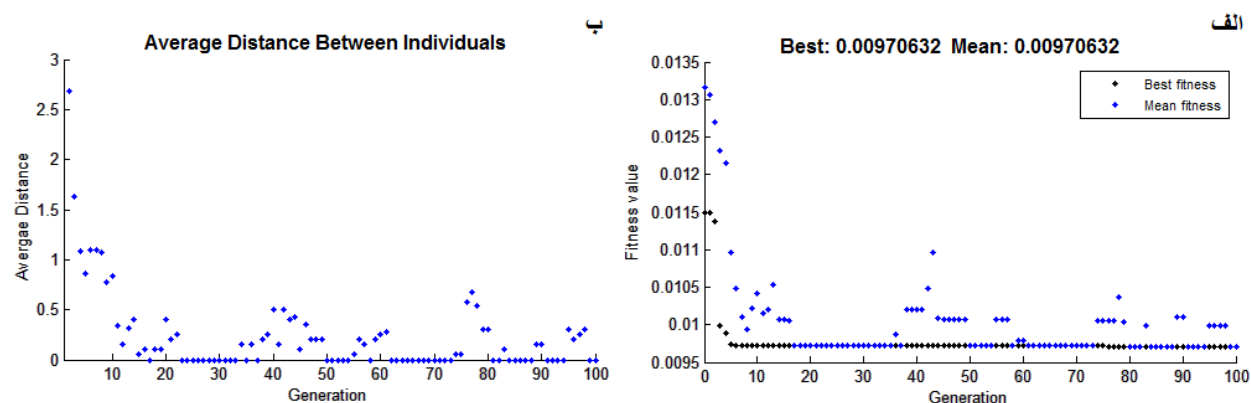
در الگوریتم GWR از هر دو هسته‌ی روابط (۴) و (۵) استفاده شد و برای بهینه‌سازی پارامتر پهنای باند نیز روش اعتبارسنجی متقابل بکارگرفته شد. ترتیب فاکتورها در تشکیل هر کروموزوم از الگوریتم ژنتیک (شکل ۱۲) به همان ترتیبی است که در جدول ۱ ارائه شده است. مقادیر پارامترهای الگوریتم ژنتیک مورد استفاده در این تحقیق نیز در جدول ۲ آمده است. در مجموع با استفاده از دو نوع هسته مختلف در GWR برای سه آتش مختلف در منطقه مورد مطالعه، ۶ حالت متفاوت خروجی داریم. پس از اجرای الگوریتم، مقدار  $R^2$  به ترتیب برای آتش‌سوزی‌های ۲۶ آبان ۱۳۸۹، ۸ فروردین ۱۳۹۰ و ۲۴ تیر ۱۳۹۰ با استفاده از هسته گوسین در GWR برابر ۰/۹۵۳۸، ۰/۹۹۹۰ و ۰/۹۹۰۳ و با استفاده از هسته مکعبی سه‌گانه در GWR برابر ۰/۹۹۳۱، ۰/۹۹۹۹ و ۰/۹۹۸۰ به دست آمد. در آتش‌سوزی ۲۶ آبان ۱۳۸۹ با استفاده از هسته گوسین، ۸ فاکتور فاصله از رودخانه‌ها، فاصله از جاده‌ها، کاربری زمین، جهت شیب، حداقل دما، متوسط دما، حداکثر سرعت باد، میزان بارندگی و با استفاده از هسته مکعبی سه‌گانه، ۸ فاکتور فاصله از مناطق مسکونی، جنس خاک، کاربری زمین، ارتفاع، مقدار شیب، جهت شیب، حداقل دما و جهت باد غالب به عنوان فاکتورهای موثر تعیین شدند. همچنین در آتش‌سوزی ۸ فروردین ۱۳۹۰ با استفاده از هسته گوسین، ۷ فاکتور فاصله از رودخانه‌ها، فاصله از



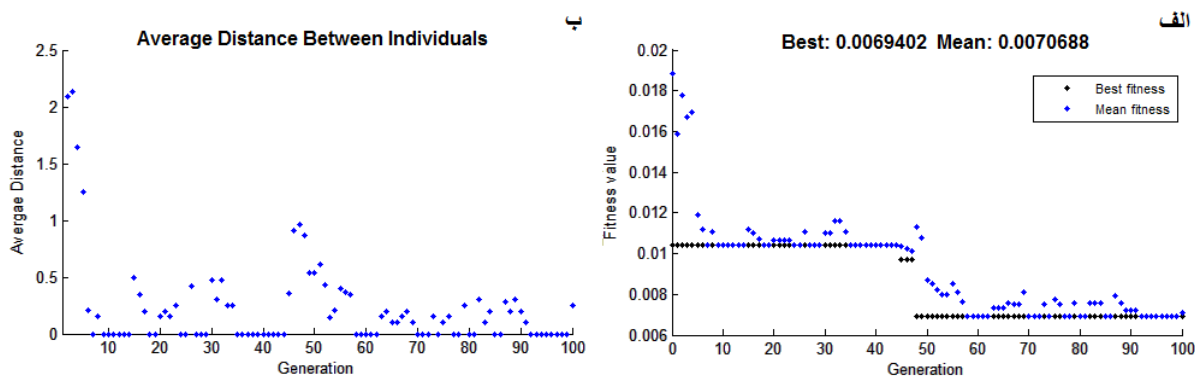
شکل ۱۶: (الف) بهترین مقدار و میانگین مقادیر تابع برازش و (ب) میانگین فاصله بین افراد هر نسل با استفاده از هسته گوسین و داده‌های ۲۶ آبان ۱۳۸۹



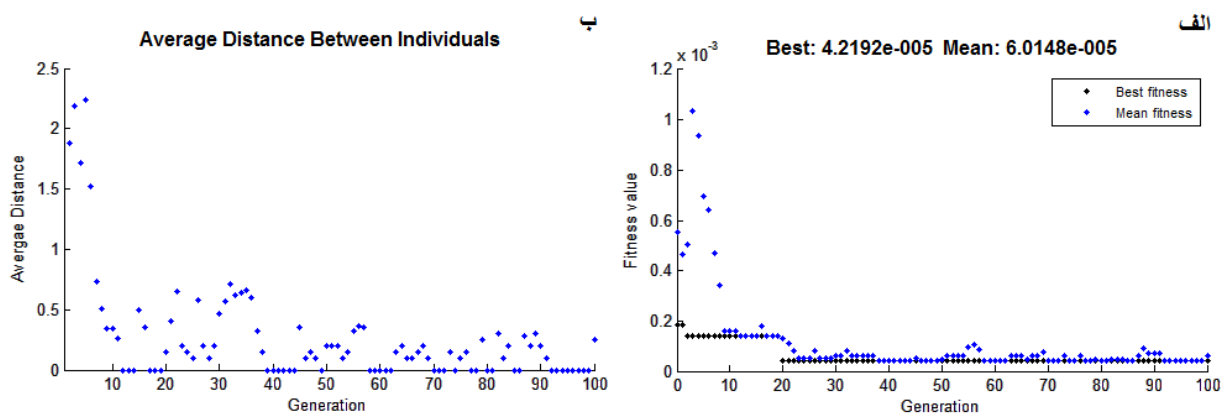
شکل ۱۷: (الف) بهترین مقدار و میانگین مقادیر تابع برازش و (ب) میانگین فاصله بین افراد هر نسل با استفاده از هسته گوسین و داده‌های ۸ فروردین ۱۳۹۰



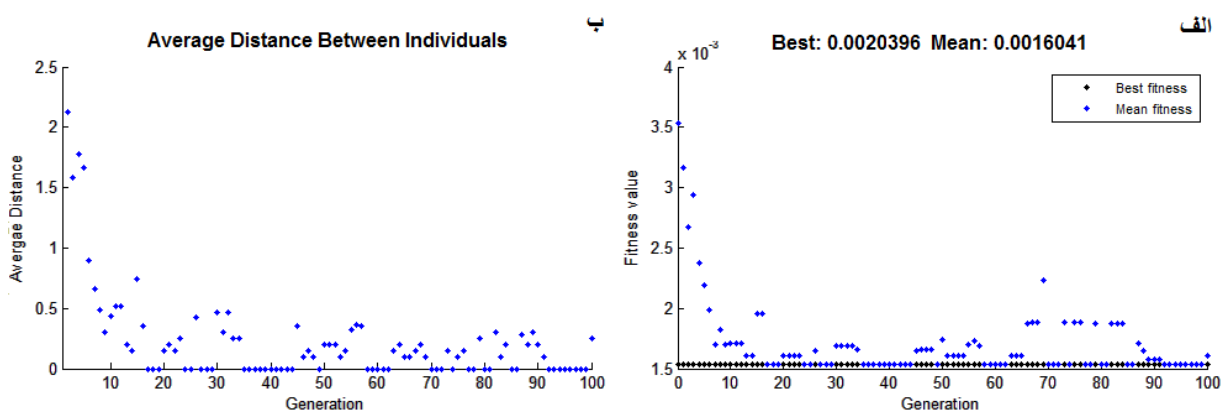
شکل ۱۸: (الف) بهترین مقدار و میانگین مقادیر تابع برازش و (ب) میانگین فاصله بین افراد هر نسل با استفاده از هسته گوسین و داده‌های ۲۴ تیر ۱۳۹۰



شکل ۱۹: (الف) بهترین مقدار و میانگین مقادیر تابع برازش و (ب) میانگین فاصله بین افراد هر نسل با استفاده از هسته مکعبی سه‌گانه و داده‌های ۲۶ آبان ۱۳۸۹



شکل ۲۰: (الف) بهترین مقدار و میانگین مقادیر تابع برازش و (ب) میانگین فاصله بین افراد هر نسل با استفاده از هسته مکعبی سه‌گانه و داده‌های ۸ فروردین ۱۳۹۰



شکل ۲۱: (الف) بهترین مقدار و میانگین مقادیر تابع برازش و (ب) میانگین فاصله بین افراد هر نسل با استفاده از هسته مکعبی سه‌گانه و داده‌های ۲۴ تیر ۱۳۹۰

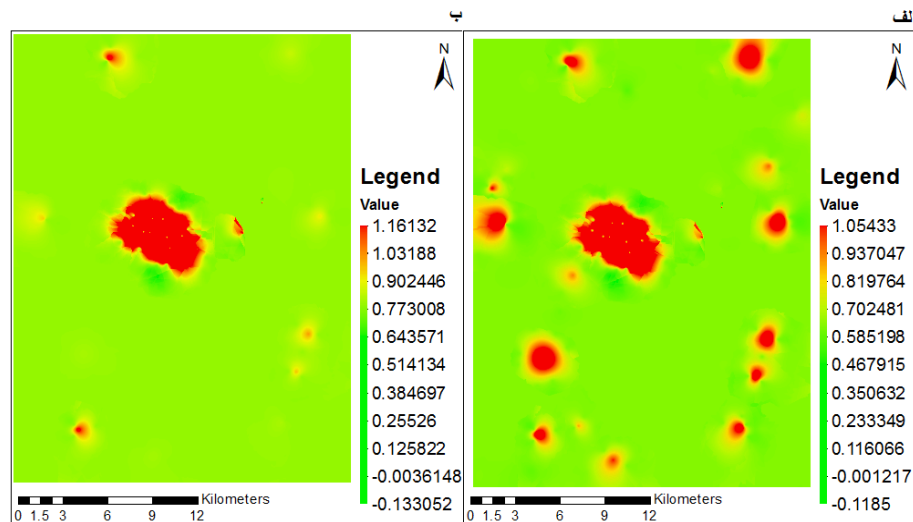
پس از اجرای الگوریتم ژنتیک و بدست آمدن ضرایب  $(\hat{\beta})$ ، مقادیر برآورد شده مشاهدات  $(\hat{y})$  و باقیمانده‌ها  $(\varepsilon)$ ؛ از روش درونیابی کریجینگ عمومی<sup>۱</sup> با مدل

<sup>۱</sup> Ordinary Kriging

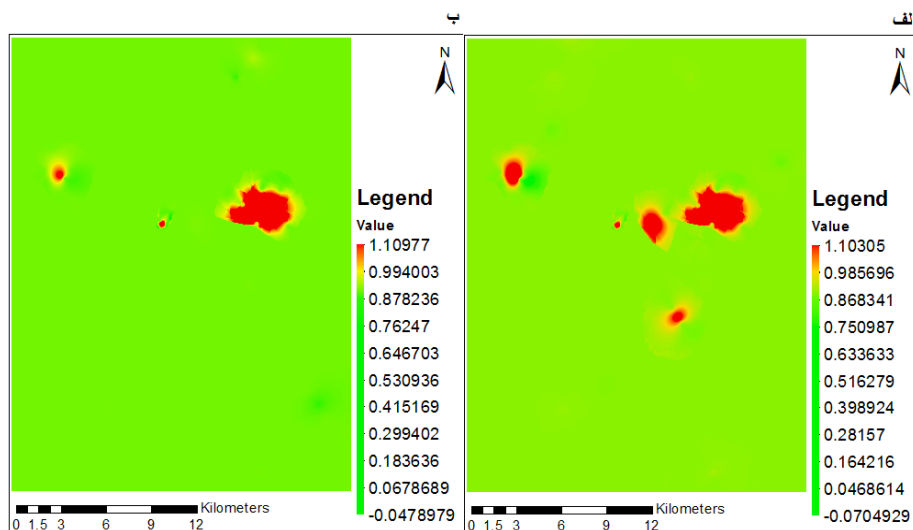


منظور محاسبه توزیع مکانی مقدار برآورد شده مشاهده آتش و باقیمانده‌ها در کل منطقه مورد مطالعه استفاده گشت که خروجی آن در شکل‌های ۲۲ تا ۲۷ ارائه شده است.

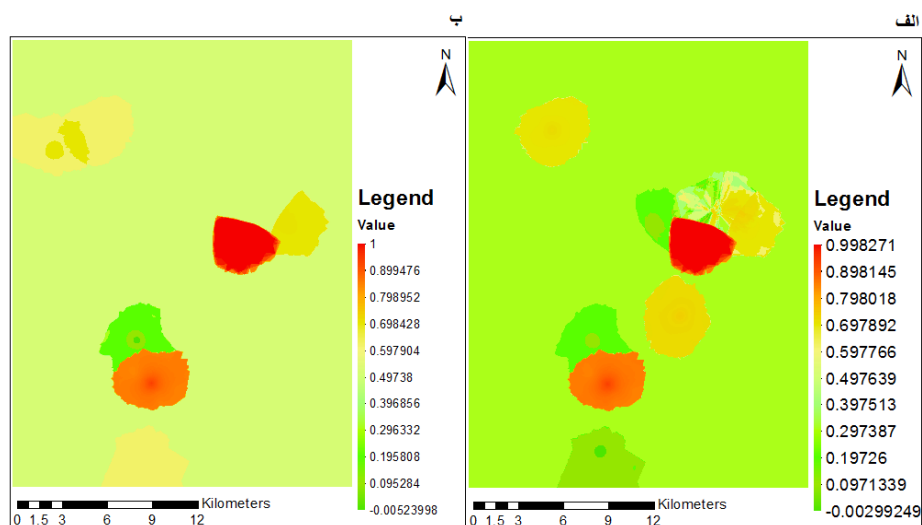
واریوگرام نمایی [۳۴] با توجه به مقادیر متوسط خطای اریب، متوسط خطای مطلق و انحراف استاندارد عمومی این مدل برای این مجموعه از داده‌ها که به ترتیب برابر با ۰/۰۰۶، ۱۵/۶۵ و ۱۰/۹۰ درصد نسبت به سایر مدل‌های واریوگرام بود و قدرت تفکیک ۳۰ متر به



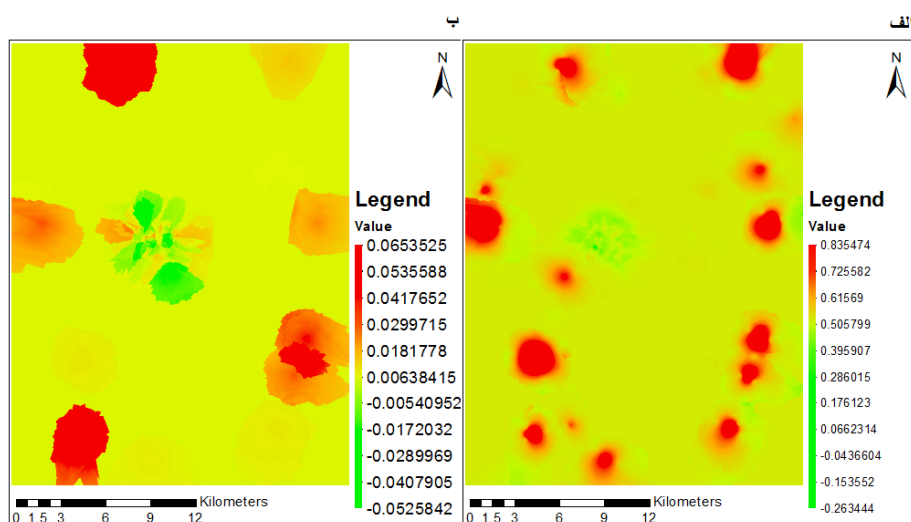
شکل ۲۲: نمایش مقادیر برآورد شده مشاهدات ۲۶ آبان ۱۳۸۹ در منطقه مورد مطالعه با استفاده از هسته (الف) گوسین و (ب) مکعبی سه‌گانه



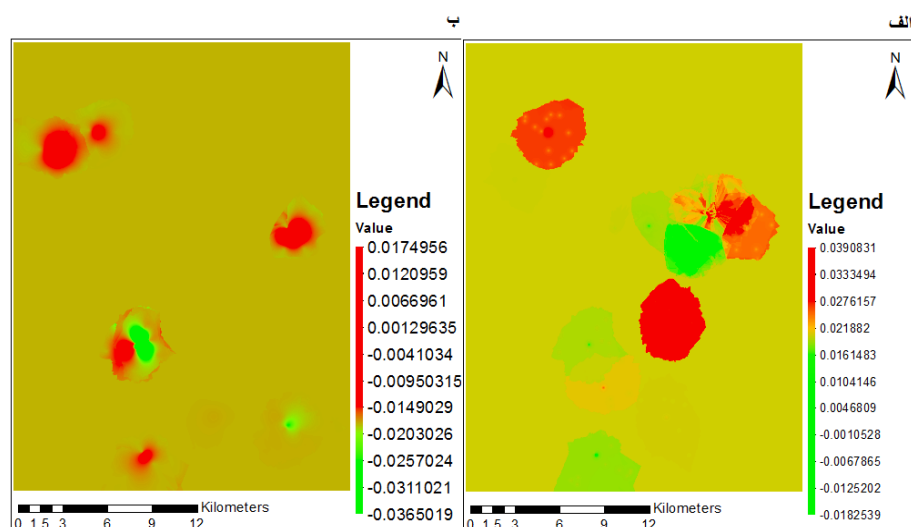
شکل ۲۳: نمایش مقادیر برآورد شده مشاهدات ۸ فروردین ۱۳۹۰ در منطقه مورد مطالعه با استفاده از هسته (الف) گوسین و (ب) مکعبی سه‌گانه



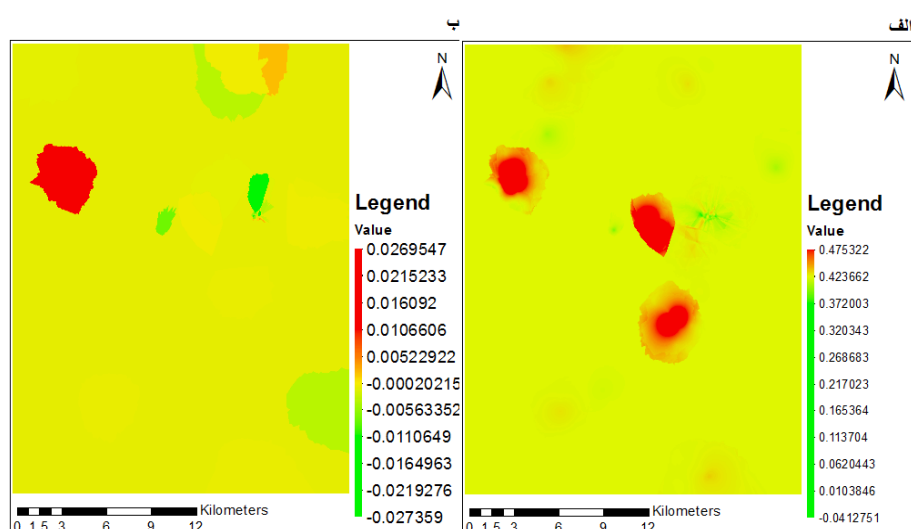
شکل ۲۴: نمایش مقادیر برآورد شده مشاهدات ۲۴ تیر ۱۳۹۰ در منطقه مورد مطالعه با استفاده از هسته (الف) گوسین و (ب) مکعبی سه‌گانه



شکل ۲۵: نمایش مقادیر باقیمانده مدل در منطقه مورد مطالعه با استفاده از داده‌های ۲۶ آبان ۱۳۸۹ و هسته (الف) گوسین و (ب) مکعبی سه‌گانه



شکل ۲۶: نمایش مقادیر باقیمانده مدل در منطقه مورد مطالعه با استفاده از داده‌های ۸ فروردین ۱۳۹۰ و هسته (الف) گوسین و مکعبی سه‌گانه (ب)



شکل ۲۷: نمایش مقادیر باقیمانده مدل در منطقه مورد مطالعه با استفاده از داده‌های ۲۴ تیر ۱۳۹۰ و هسته (الف) گوسین و (ب) مکعبی سه‌گانه

تشخیص داده نشدند با علامت (-) مشخص شده‌اند و شماره ۰ نیز مربوط به مقدار ثابت رگرسیون است. مقادیر RMSE، NRMSE و شاخص موران نیز در جدول ۴ آمده است.

انحراف استاندارد ضرایب رگرسیون در جدول ۳ ارائه شده است. لازم به ذکر است که ترتیب ضرایب در جدول ۳ با ترتیب فاکتورها در جدول ۱ یکی است. ضرایب مربوط به فاکتورهایی که موثر

جدول ۳: انحراف استاندارد ضرایب رگرسیون حاصل از GWR در دو حالت استفاده از هسته (الف) گوسین و (ب)

مکعبی سه‌گانه

(ب) مکعبی سه‌گانه			(الف) گوسین			شماره
۲۴ تیر ۱۳۹۰	۸ فروردین ۱۳۹۰	۲۶ آبان ۱۳۸۹	۲۴ تیر ۱۳۹۰	۸ فروردین ۱۳۹۰	۲۶ آبان ۱۳۸۹	
۰/۱۸۵۹۴۶	۰/۰۴۲۹۸۳	۵/۰۵۷۸۵۸	۰/۰۹۱۳۵۸	۰/۱۱۱۳۵۸	۱/۷۳۳۲۰۶	۰
-	-	-	-	۰/۰۰۰۱۸۴	۰/۰۰۰۰۸۶	۱
-	۰/۰۰۰۳۰۵	-	-	-	۰/۰۰۰۱۱۴	۲
۰/۰۰۱۱۷۰	-	۰/۰۰۵۲۸۴	۰/۰۰۰۵۵۷	۰/۰۰۰۱۴۹	-	۳
۱۲/۴۵۷۸۴۱	۲/۸۴۰۶۴۹	۲/۸۹۵۷۸۲	-	-	-	۴
-	۰/۰۱۹۶۶۸	۶/۸۱۴۷۸۸	۰/۰۲۷۵۰۳	۰/۰۶۸۱۹۴	۰/۰۵۶۱۹۸	۵
۰/۰۰۴۸۵۲	۰/۰۰۰۱۵۷	۰/۰۱۱۰۴۷	۰/۰۰۰۹۹۵	-	-	۶
۰/۰۰۲۷۴۵	-	۰/۰۲۴۳۴۲	-	۰/۰۰۱۰۵۹	-	۷
۰/۰۰۰۴۵۴	-	۰/۰۰۷۳۶۵	-	۰/۰۰۰۵۱۰	۰/۰۰۰۱۶۹	۸
-	-	-	۱/۶۶۳۸۲۶	-	-	۹
۳۱/۸۷۴۱۴۳	-	۴۸/۹۹۶۴۰۱	۴/۲۴۴۶۸۷	-	۲۶/۲۸۴۷۶۱	۱۰
-	۰/۶۸۹۷۱۰	-	۲/۷۴۱۰۷۳	۱/۷۶۰۰۷۸	۱۵/۱۰۳۹۸۰	۱۱
-	-	۱/۷۲۰۷۷۷	۰/۵۲۶۳۱۰	-	-	۱۲
۱۵/۵۷۹۰۹۴	-	-	-	۴/۱۳۸۳۵۸	۰/۷۴۶۹۵۴	۱۳
-	۱۹/۰۴۱۹۳۰	-	-	-	۱۸/۷۲۷۱۹۲	۱۴

جدول ۴: مقدار RMSE، NRMSE و شاخص موران باقیمانده‌های مدل

(ب) مکعبی سه‌گانه			(الف) گوسین			
۲۴ تیر ۱۳۹۰	۸ فروردین ۱۳۹۰	۲۶ آبان ۱۳۸۹	۲۴ تیر ۱۳۹۰	۸ فروردین ۱۳۹۰	۲۶ آبان ۱۳۸۹	
۰/۰۰۹۰۶۲	۰/۰۰۰۷۸۶	۰/۰۲۰۹۳۷	۰/۰۳۹۴۰۸	۰/۰۰۶۵۶۶	۰/۰۹۵۹۳۶	RMSE
۰/۰۲۲۷۴۱	۰/۰۰۳۷۵۰	۰/۰۴۸۴۳۳	۰/۰۹۹۴۴۷	۰/۰۳۱۵۹۷	۰/۲۲۹۳۲۵	NRMSE
۰/۶۳۲۱۴۵	-۰/۴۸۲۱۵۹	-۰/۵۳۶۴۲۱	۰/۷۶۸۹۸۱	۰/۶۳۷۵۳۱	۰/۶۹۳۰۲۱	Moran's

### ۳-۱- ارزیابی نتایج

با استفاده از هسته گوسین هر دو برابر با ۰/۰۰۰۹۸۲ و با استفاده از هسته مکعبی سه‌گانه به ترتیب برابر با ۰/۰۰۰۰۴۲ و ۰/۰۰۰۰۶۰ و همچنین برای آتش‌سوزی ۲۴ تیر ۱۳۹۰ با استفاده از هسته گوسین هر دو برابر با ۰/۰۰۹۷۰۶ و با استفاده از هسته مکعبی سه‌گانه به ترتیب برابر با ۰/۰۰۲۰۴۰ و ۰/۰۰۱۶۰۴ به دست آمده است. این نتایج نشان می‌دهد که توانایی هسته مکعبی سه‌گانه در مدل‌سازی روابط بین

همان‌طور که در شکل‌های ۱۶ تا ۲۱ نشان داده شده است، بهترین مقدار تابع برازش و میانگین مقادیر آن برای آتش‌سوزی ۲۶ آبان ۱۳۸۹ با استفاده از هسته گوسین به ترتیب برابر با ۰/۰۴۶۱۶۳ و ۰/۰۴۶۹۰۰ و با استفاده از هسته مکعبی سه‌گانه به ترتیب برابر با ۰/۰۰۶۹۴۰ و ۰/۰۰۷۰۶۸ برای آتش‌سوزی ۸ فروردین ۱۳۹۰



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.

برآورد آتش که در شکل‌های ۲۲ تا ۲۴ آمده در حالت استفاده از هسته مکعبی سه‌گانه نسبت به هسته گوسین به آتش واقعی نزدیک‌تر است اما همچنان به دلیل عدم دسترسی به داده‌های کافی و مناسب دارای خطاهایی نیز می‌باشد و در همه موارد برآورد آتش نسبت به آتش واقعی مساحت بزرگتری را شامل می‌شود. باقیمانده‌های مدل نیز که در شکل‌های ۲۵ تا ۲۷ نشان داده شده‌اند، در حالت استفاده از هسته مکعبی سه‌گانه نسبت به هسته گوسین کوچک‌تر هستند. جدول ۳ نشان‌دهنده نایستایی مکانی نسبتاً بالای مقدار ثابت رگرسیون و ضرایب فاکتورهای ۴، ۵، ۹، ۱۰، ۱۱، ۱۲ و ۱۳ به خصوص برای داده‌های ۲۶ آبان ۱۳۸۹ می‌باشد. در سایر موارد میزان این نایستایی بسیار پایین است. همچنین جدول ۴ نشان می‌دهد که مقدار RMSE و NRMSE باقیمانده‌های مدل در هر دو حالت استفاده از هسته گوسین و مکعبی سه‌گانه نزدیک به صفر است؛ هرچند در حالت استفاده از هسته مکعبی سه‌گانه باقیمانده‌ها بسیار کوچک‌تر هستند.

همچنین در جدول ۴ شاخص موران در هر دو حالت استفاده از هسته گوسین و مکعبی سه‌گانه آمده است که نشان‌دهنده خودهمبستگی مکانی نسبتاً بالای بین باقی‌مانده‌هاست که این موضوع نشان‌دهنده ناکافی بودن و نقص در داده‌های ورودی است. با مقایسه شاخص موران در دو حالت نیز می‌توان به برتری استفاده از هسته مکعبی سه‌گانه نسبت به هسته گوسین پی برد. در مجموع می‌توان گفت با توجه به تمامی ملاک‌ها استفاده از هسته مکعبی سه‌گانه به نتایج بهتری انجامید.

آتش‌سوزی و فاکتورهای موجود در جدول ۱ برای منطقه مورد مطالعه بیشتر است. همچنین میانگین فاصله بین افراد هر نسل یا اصطلاحاً تنوع<sup>۱</sup> را می‌بینیم که در همه حالت‌ها تقریباً مشابه است. به طور کلی در تمام حالت‌ها این فاصله کاهش می‌یابد اما نوسان‌های زیادی نیز در آن‌ها دیده می‌شود.

فاکتور کاربری زمین در ۵ حالت، فاکتورهای فاصله از مناطق مسکونی، ارتفاع، جهت شیب، حداقل دما و متوسط دما در ۴ حالت، فاکتورهای جنس خاک، مقدار شیب و حداکثر سرعت باد در ۳ حالت، فاکتورهای فاصله از رودخانه‌ها، فاصله از جاده‌ها، جهت باد غالب و میزان بارندگی در ۲ حالت و فاکتور حداکثر دما تنها در ۱ حالت موثر شناخته شدند. همچنین طبق آمارهای چند سال اخیر آتش‌سوزی‌های جنگل گلستان، بیشتر آن‌ها توسط انسان اتفاق افتادند بنابراین اهمیت بالای فاکتورهای کاربری زمین و فاصله از مناطق مسکونی در این تحقیق با آمارها نیز مطابقت می‌کند. حضور عوامل بیوفیزیکی و انسانی در فاکتورهای بهینه نشان می‌دهد که در مدیریت آتش‌سوزی جنگل‌ها باید به هر دو دسته این عوامل توجه نمود. البته باید توجه داشت که در این تحقیق تنها ۳ فاکتور انسانی وجود داشت که از این میان ۲ فاکتور در حداقل ۴ حالت از خروجی‌ها مهم تشخیص داده شده‌اند که این با فرض اولیه تحقیق همخوانی دارد ولی 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. در واقع 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

<sup>22</sup> diversity

#### ۴- نتیجه‌گیری

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

در این تحقیق سعی شد فاکتورهای موثر بر آتش‌سوزی‌های جنگل گلستان را شناسایی نمائیم و برای این منظور از روش رگرسیون وزن‌دار جغرافیایی (GWR) در ترکیب با الگوریتم ژنتیک استفاده شد. روش مناسبی برای مسائل رگرسیون مکانی است زیرا دو ویژگی خاص داده‌های مکانی یعنی خودهمبستگی و ناپایداری مکانی را در حل مسئله در نظر می‌گیرد. در این تحقیق، از دو حالت وزن‌دهی گوسین و مکعبی سه‌گانه در GWR استفاده شد. نتایج این تحقیق نشان داد که هر دو دسته فاکتورهای بیوفیزیکی و انسانی تأثیر بسزایی در آتش‌سوزی‌های منطقه مورد مطالعه

داشتند. از فاکتورهای بیوفیزیکی ارتفاع، جهت شیب، حداقل دما، متوسط دما و از فاکتورهای انسانی کاربری زمین و فاصله از مناطق مسکونی در بیشتر حالت‌ها موثر شناخته شدند و در مجموع با استفاده از هسته مکعبی سه‌گانه برای وزن‌دهی در GWR نتایج دقیق‌تر و مناسب‌تری به دست آمد.

در تحقیقات آینده پیشنهاد می‌شود تعداد فاکتورها افزایش یابد. در این زمینه می‌توان به فاکتورهای انسانی مانند سطح تحصیلی، فرهنگی و اقتصادی مردم، تعداد گردشگر، قوانین و سیاست‌های مدیریت جنگل و مدیریت بحران اشاره نمود. اگرچه جمع‌آوری این اطلاعات از چالش‌های پیش روست. از طرف دیگر فاکتورهای بیوفیزیکی دیگری نیز می‌توان به مسئله افزودن از جمله نوع و تراکم پوشش گیاهی و ویژگی‌های سوختی آن‌ها، رطوبت هوا، ساعات آفتابی و درصد بخار آب که متأسفانه به دلیل نبود 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 سنجش 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.

#### مراجع

- [1] Y. Bergeron, S. Gauthier, M. Flannigan, and V. Kafka, "Fire regimes at the transition between mixed wood and coniferous boreal forest in Northwestern Quebec", *Ecology*, Vol.85, No.7, PP. 1916–1932, 2004.
- [2] P. F. Hessburg, J. K. Agee, and J. F. Franklin, "Dry forests and wildland fires of the inland Northwest USA: Contrasting the landscape ecology of the pre-settlement and modern eras", *For. Ecol. Manage.* Vol.211, No.1–2, PP. 117–139, 2005.
- [3] J. Martinez-Fernandez, E. Chuvieco and N. Koutsias, "Modeling long-term fire occurrence factors in Spain by accounting for local variations with geographically weighted regression", *Natural Hazards and Earth System Sciences*, Vol.13, PP. 311–327, 2013.

- [4] B. G. Dickson, J. W. Prather, Y. Xu, H. M. Hampton, E. N. Aumack and T. D. Sisk, "Mapping the probability of large fire occurrence in northern Arizona, USA", *Landscape Ecology*, Vol.21, PP. 747–761, 2006.
- [5] D. C. Odion, E. J. Frost, J. R. Strittholt, H. Jiang, D. A. Della-Salla and M. A. Moritz, "Patterns of fire severity and forest conditions in the western Klamath Mountains, north-western California" *Conservation Biology*, Vol.18, PP. 927–936, 2004.
- [6] J. P. Prestemon, J. M. Pye, D. T. Butry, T. P. Holmes and D. E. Mercer, "Understanding broad scale wildfire risks in a human-dominated landscape", *For. Sci.* Vol.48, PP. 685– 693, 2002.
- [7] D. E. Mercer and J. P. Prestemon, "Comparing production function models for wildfire risk analysis in the wildland-urban interface", *For. Policy Economics*, Vol.7, No.5, PP. 782–795, 2005.
- [8] M. A. Moritz, J. E. Keeley, E. A. Johnson and A. A. Schaffner, "Testing a basic assumption of shrubland fire management: How important is fuel age?", *Frontiers Ecol. Environm.* Vol.2, PP. 67–72, 2004.
- [9] J. Yang, H. S. He, S. R. Shifley and E. J. Gustafson, "Spatial Patterns of Modern Period Human-Caused Fire Occurrence in the Missouri Ozark Highlands", *Forest Science*, Vol.53, No.1, PP. 1-15, 2007.
- [10] C. M. Countryman, "The fire environment concept", USDA Forest Service, Pacific Southwest Forest and Range Experiment Station, General Technical Report PSW-7. Berkeley, CA, 1972.
- [11] R. A. Bradstock, A. M. Gill, B. J. Kenny and J. Scott, "Bushfire risk at the urban interface estimated from historical weather records: consequences for the use of prescribed fire in the Sydney region of south-eastern Australia", *Journal of Environmental Management*, Vol.52, PP. 259–271, 1998.
- [12] J. S. Fried, G. Winter and J. K. Gilles, "Assessing the benefits of reducing fire risk in the wildland–urban in-terface: a contingent valuation approach", *International Journal of Wildland Fire*, Vol.9, PP. 9–20, 1999.
- [13] C. Diaz-Avalos, D. L. Peterson, E. Alvarado, S. A. Ferguson and J. E. Besag, "Space–time modeling of lightning-caused ignitions in the Blue Mountains, Oregon", *Canadian Journal of Forest Research*, Vol.31, PP. 1579–1593, 2001.
- [14] R. E. Burgan and R. C. Rothermel, "BEHAVE: fire prediction and fuel modeling system–fuel subsystem", USDA Forest Service, Intermountain Research Station, General Technical Report INT-167. Ogden, UT, 1984.
- [15] Forestry Canada Fire Danger Group Development and structure of the Canadian Forest Fire Behaviour Prediction System, Forestry Canada, Science and Sustainable Development Directorate Report ST-X-3. Ottawa, ON, 1992.
- [16] M. Larjavaara, J. Pennanen and T. J. Tuomi, "Lightning that ignites forest fires in Finland", *Agricultural and Forest Meteorology*, Vol.132, PP. 171–180, 2005.
- [17] B. M. Wotton and D. L. Martell, "A lightning fire occurrence model for Ontario", *Canadian Journal of Forest Research*, Vol.35, PP. 1389–1401, 2005.
- [18] K. L. Pew and C. P. S. Larsen, "GIS analysis of spatial and temporal patterns of human-caused wildfires in the temperate rainforest of Vancouver Island, Canada", *Forest Ecology and Management*, Vol.140, PP. 1–18, 2001.
- [19] Y. H. Chou, "Spatial autocorrelation and weighting functions in the distribution of wildland fires", *International Journal of Wildland Fire*, Vol.2, PP. 169–176, 1992.
- [20] Y. H. Chou, R. A. Minnich and R. A. Chase, "Mapping probability of fire occurrence in San Jacinto Mountains, California, USA", *Environmental Management*, Vol.17, PP.

- 129–140, 1993.
- [21] A. D. Syphard, V. C. Radeloff, J. E. Keeley, T. J. Hawbaker, M. K. Clayton, S. I. Stewart and R. B. Hammer, "Human influence on California fire regimes", *Ecological Applications*, Vol.17, PP. 1388–1402, 2007.
- [22] T. D. A. Rodriguez, M. H. Ramirez and J. H. Tchikoue, "Factors affecting the accident rate of forest fires", *Ciencia Forestal en Mexico*, Vol.33, No.104, PP. 38–57, 2008.
- [23] C. R. Roman and V. M. J. Martinez, "Effectiveness of protected areas in mitigating fire within their boundaries: Case study of Chiapas, Mexico", *Conservation Biology*, Vol.20, No.4, PP. 1074–1086, 2006.
- [24] A. D. Syphard, V. C. Radeloff, N. S. Keuler, R. S. Taylor, T. J. Hawbaker, S. I. Stewart and M. K. Clayton, "Predicting spatial patterns of fire on a southern California landscape", *International Journal of Wildland Fire*, Vol.17, PP. 602–613, 2008.
- [25] R. Romero-Calcerrada, C. J. Novillo, J. D. A. Millington and I. Gomez-Jimenez, "GIS analysis of spatial patterns of human-caused wildfire ignition risk in the SW of Madrid (Central Spain)", *Landscape Ecol.* Vol.23, PP. 341–354, 2008.
- [26] E. Erten, V. Kurgun, and N. Musaolu, "Forest fire risk zone mapping from satellite imagery and risk: A recent case study from Mt. Carmel (Israel)", *Forest Ecology and Management*, Vol.262, PP. 1184–1188, 2005.
- [27] F. D. Y. Avila, G. M. Pompa and P. E. Vargas, "Spatial analysis of fire occurrence in the Durango state. *Revista Chapingo*", *Serie Ciencias Forestales*, Vol.16, No.2, 2010.
- [28] C. Brunson, S. Fotheringham and M. Charlton, "Geographically weighted regression – modelling spatial non-stationarity", *The Statistician*, Vol.47, No.3, PP. 431–443, 1998.
- [29] F. D. Y. Avila, M. Pompa-Garcia, X. Antonio-Nemiga, D.A. Rodriguez-Trejo, E. Vargas-Perez and J. Santillan-Perez, "Driving Factors for Forest Fire Occurrence in Durango State of Mexico: A Geospatial Perspective", *Chin. Geogra. Sci.* Vol.20, No.6, PP. 491–497, 2010.
- [30] A. C. L. Sa, J. M. C. Pereira, M. E. Charlton, B. Mota, P. M. Barbosa and A. S. Fotheringham, "The pyrogeography of sub-Saharan Africa: a study of the spatial non-stationarity of fire–environment relationships using GWR", *J. Geogr. Syst.* Vol.13, PP. 227–248, 2011.
- [31] W. R. Tobler, "A computer movie simulating urban growth in the Detroit region", *Economic Geography*, Vol.46, No.2, PP. 234–24, 1970.
- [32] D. P. McMillen and J. F. McDonald, "Locally weighted maximum likelihood estimation: Monte Carlo evidence and an application", presented at the Regional Science Association International meetings, Santa Fe, NM, 1998.
- [33] M. Charlton and A. S. Fotheringham, "Geographically Weighted Regression", White Paper. Kildare, Ireland: National Centre for Geocomputation, National University of Ireland, Maynooth, 17, 2009.
- [34] S. Shekhar, and H. Xiong. "Encyclopedia of GIS", Springer Science & Business Media, 2008.
- [35] A. Cliff and K. Ord, "Testing for spatial autocorrelation among regression residuals", *Geographical Analysis* Vol.4, No.3, PP. 267–284, 1972.
- [36] M. Hasanlou and F. Samadzadegan, "ICA/PCA base genetically band selection for classification of Hyperspectral images", presented at the 31st Asian Conference on Remote Sensing, announcer, 2010.
- [37] P. Dale. "Mathematical Techniques in GIS, Second Edition", CRC Press, 2014.





## **Determining Effective Factors on Forest Fire Using the Compound of Geographically Weighted Regression and Genetic Algorithm, a Case Study: Golestan, Iran**

Amin Raei<sup>1</sup>, Parham Pahlavani<sup>\*2</sup>, Mahdi Hasanlou<sup>2</sup>

1- M.Sc student of GIS in Department of Surveying and Geomatics Engineering, College of Engineering, University of Tehran

2- Assistant professor in Department of Surveying and Geomatics Engineering, College of Engineering, University of Tehran

### **Abstract**

Determining the effective factors on fire is so important, because the plenty areas of forests around the world are destroyed every year by fire. It helps us to identify most dangerous locations and times in forest fire. Hence, we can prevent many of driving factors of forest fire by law enforcement, efficient forest management policies and more supervision. In the current study, we identified the impressive factors on the fire in Golestan forest using the compound of Geographically Weighted Regression (GWR) method and Genetic Algorithm that is suitable for the spatial regression problem, because it obtains the effective factors considering the autocorrelation and non-stationarity properties of spatial data. In this study, three different fire areas as well as two kernels of Gaussian and Tricube for weighting of GWR were used that for these three fire areas resulted to  $R^2=0.9538$ ,  $R^2=0.9990$ , and  $R^2=0.9903$  for Gaussian kernel and  $R^2=0.9931$ ,  $R^2=0.9999$ , and  $R^2=0.9980$  for Tricube kernel, respectively. This research shows that both of the biophysical and anthropogenic factors have significant effects on forest fire in our study areas. In biophysical factors, the elevation, the aspect, the minimum and mean temperature and in anthropogenic factors, the landuse and the distance from the residential areas were identified as the most impressive factors. Weighting by Tricube kernel concluded to more precise results.

**Key words:** Forest Fire, Geographically Weighted Regression, Genetic Algorithm, Golestan Forest