



جغرافیا و روابط انسانی، زمستان ۱۴۰۳، دوره ۷، شماره ۴، صص ۱۶۰-۱۳۷

تحلیل مکانی و مدل‌سازی بروز سرطان پوست در ایران با استفاده از GIS

بهمن محمدی^{۱*} محمد رئوف حیدری فر^۲

۱- دانشگاه تهران، دانشکده جغرافیا، گروه سنجش از دور و سیستم اطلاعات جغرافیایی، تهران، ایران،

muhammadi.93@ut.ac.ir

۲- دانشیار جغرافیای سیاسی، دانشگاه پیام نور، تهران، ایران

تاریخ پذیرش: ۱۴۰۳/۱۱/۳۰

تاریخ دریافت: ۱۴۰۳/۰۹/۲۴

چکیده

سرطان پوست به‌عنوان یکی از شایع‌ترین انواع سرطان در ایران، در سال‌های اخیر روند روبه‌رشدی داشته و به دلیل پیامدهای جسمی، روانی و هزینه‌های بالای درمان، به یکی از دغدغه‌های اصلی نظام سلامت تبدیل شده است. مطالعات مختلف نشان داده‌اند که عوامل طبیعی و اقلیمی نقش مهمی در بروز این بیماری ایفا می‌کنند. هدف اصلی مکانی، اقلیمی و به‌کارگیری دو روش پیشرفته مدل‌سازی شامل رگرسیون وزندار جغرافیایی (GWR) و شبکه عصبی پرسپترون چندلایه (MLP) است.

ابتدا داده‌های مربوط به بروز سرطان پوست، پارامترهای اقلیمی و جمعیتی گردآوری شد. پس از تصحیح، فازی‌سازی و پیاده‌سازی تحلیل‌های مکانی با استفاده از سامانه اطلاعات جغرافیایی (GIS)، مدل‌سازی با دو رویکرد MLP و GWR انجام گرفت. نتایج حاصل از اعتبارسنجی مدل‌ها نشان داد که مدل MLP با ضریب همبستگی ۰.۷۴ نسبت به مدل GWR با ضریب ۰.۴۹ عملکرد دقیق‌تری در پیش‌بینی نرخ بروز سرطان پوست دارد. تحلیل حساسیت نیز مشخص کرد که مجموع بارش، ساعات آفتابی، میانگین ارتفاع، تعداد روزهای گردوغبار و تعداد روزهای برف و تگرگ مهم‌ترین عوامل مؤثر بر بروز این بیماری هستند.

یافته‌های پژوهش حاضر می‌تواند در راستای سیاست‌گذاری‌های بهداشتی، شناسایی مناطق پرخطر، اولویت‌بندی مداخلات پیشگیرانه و افزایش آگاهی عمومی کاربرد داشته باشند.

واژگان کلیدی: سرطان پوست، مدل‌سازی مکانی، GIS، شبکه عصبی پرسپترون چندلایه (MLP)، رگرسیون وزندار جغرافیایی (GWR)، عوامل اقلیمی.

۱ مقدمه

در دهه‌های اخیر، پیشرفت‌های علمی و فناوری، توانایی انسان در درک و تعامل با محیط طبیعی را به شکل چشمگیری افزایش داده است. با این حال، تأثیرات منفی محیط بر سلامت انسان همچنان باقی مانده و بیماری‌های ناشی از عوامل محیطی، به ویژه در ارتباط مستقیم و بلندمدت انسان با طبیعت، یکی از دغدغه‌های اصلی حوزه سلامت عمومی به شمار می‌رود. شناسایی عوامل مؤثر بر بروز بیماری‌ها، تحلیل فضایی و مکانی آن‌ها، کشف الگوهای انتشار، و مدل‌سازی دقیق شیوع بیماری‌ها، از نیازهای اساسی در بهداشت عمومی و سیاست‌گذاری سلامت هستند (بیاتانی و صادقی، ۱۳۹۱).

بیماری‌ها اغلب دارای ویژگی‌های مکانی هستند و بررسی آن‌ها نیازمند ابزارهای تحلیل فضایی است (آذرمهر و همکاران، ۱۳۸۹). در این میان، سیستم‌های اطلاعات جغرافیایی (GIS) به عنوان ابزاری توانمند در جمع‌آوری، ذخیره‌سازی، تحلیل و نمایش داده‌های مکانی، نقش مهمی در تحلیل الگوهای بیماری و پیش‌بینی مناطق در معرض خطر ایفا می‌کنند. استفاده از GIS در پایش بیماری‌ها، تحلیل توزیع جغرافیایی آن‌ها، تهیه نقشه‌های خطر، مکان‌یابی بهینه امکانات درمانی، و همچنین مدل‌سازی شیوع بیماری بر اساس عوامل طبیعی و اقلیمی به اثبات رسیده است (زارع و همکاران، ۱۳۸۵؛ ملالو، ۱۳۹۲؛ آهنگرکانی، ۱۳۹۵).

سرطان پوست، به عنوان شایع‌ترین نوع سرطان در ایران طی سال‌های ۱۳۸۲ تا ۱۳۸۸ در میان کل جمعیت گزارش شده است (گزارش کشوری ثبت موارد سرطان سال ۱۳۸۲-۱۳۸۹، ۱۳۸۹)، و از جمله بیماری‌هایی است که عوامل محیطی و مکانی در بروز آن نقش بسزایی دارند. (Wallingford, 2014) هزینه‌های بالای درمان، همراه با پیامدهای اجتماعی و روانی آن، ضرورت پرداختن به راهکارهای پیشگیرانه و شناسایی مناطق پرخطر را دوچندان می‌کند. در این راستا، پژوهش حاضر با هدف تحلیل مکانی سرطان پوست در سطح شهرستان‌های کشور ایران، با بهره‌گیری از داده‌های اقلیمی و استفاده از GIS و مدل‌های پیشرفته‌ای چون شبکه عصبی مصنوعی پرسپترون چندلایه (MLP) و رگرسیون وزندار جغرافیایی (GWR) انجام شده است.

در این تحقیق، تلاش شده است تا علاوه بر تهیه نقشه‌های توزیع جغرافیایی سرطان پوست، عوامل اقلیمی مؤثر بر بروز این بیماری شناسایی و اولویت‌بندی شوند. همچنین کارایی دو مدل GWR و MLP در مدل‌سازی نرخ بروز سرطان پوست مورد مقایسه قرار گرفته است تا مناسب‌ترین روش برای تحلیل دقیق‌تر تعیین شود. بر این اساس، سؤالات کلیدی تحقیق شامل موارد زیر هستند:

۱. آیا می‌توان بر اساس عوامل طبیعی و اقلیمی، نرخ بروز سرطان پوست را مدل‌سازی کرد؟

۲. اولویت‌بندی این عوامل به چه صورت است؟

۳. کدام یک از دو مدل GWR و MLP در مدل‌سازی دقیق‌تر موفق‌تر عمل می‌کنند؟

در راستای پاسخ به این سؤالات، فرضیات تحقیق بر پایه این مفروضات شکل گرفته است که:

- با استفاده از مدل‌های مذکور می‌توان نرخ بروز بیماری را مدل‌سازی کرد.
- مدل شبکه عصبی MLP در مقایسه با مدل GWR عملکرد بهتری دارد.

- عوامل محیطی، به ویژه ساعات آفتابی، در بروز سرطان پوست نقش پررنگی ایفا می‌کنند.

پژوهش حاضر از منظر نوآوری نیز حائز اهمیت است. نخست، برخلاف تحقیقات پیشین که عمدتاً بر تحلیل‌های آماری کلاسیک تکیه داشتند، در این تحقیق از مدل‌سازی پیشرفته با در نظر گرفتن متغیرهای اقلیمی استفاده شده است. دوم، با بهره‌گیری از روش‌های دقیق درون‌یابی، نقشه‌های با دقت بالای عوامل اقلیمی تولید گردیده است. سوم، برای نخستین بار، صحت‌سنجی داده‌های مکانی بیماران و اصلاح خطاهای آدرس‌دهی انجام شده که این امر موجب افزایش اعتبار یافته‌ها شده است و در نهایت، روش ارائه‌شده در این تحقیق، روشی سریع و کم‌هزینه برای تولید نقشه‌های اقلیمی مرتبط با سلامت است که می‌تواند در سیاست‌گذاری‌های بهداشتی و آگاه‌سازی عمومی نسبت به عوامل خطر ساز مؤثر واقع شود.

۲ تاریخچه‌ای از موضوع تحقیق

مطالعات اولیه نشان داده‌اند که نقشه‌ها ابزاری مفید برای تحلیل داده‌های مکانی در بهداشت عمومی هستند. در سال ۱۸۴۰، کووان در گلاسگو رابطه بین ازدحام جمعیت و شیوع تب زرد را بررسی کرد. (Burrough, 1986) اسنو در سال ۱۸۵۴، گسترش مکانی مرگ‌ومیر ناشی از وبا در لندن را با آب‌های آلوده مرتبط دانست. (Snow, 1856) در ایران، معصومی (۱۳۸۴) ارتباط بین سرطان و عوامل ایجاد کننده آن را با استفاده از GIS در ۱۸ استان بررسی کرد. رجبی و همکاران (۲۰۱۲) از روش‌های تصمیم‌گیری چند معیاره و GIS برای مدل‌سازی لیشمانیوز احشایی در آذربایجان شرقی استفاده کردند. حسن و همکاران (۱۹۹۹) نشان دادند که وقوع سالک در شمال شرق مصر وابسته به پوشش گیاهی و نوع خاک است. ملالو (۱۳۹۲) نیز در استان گلستان نشان داد که توپوگرافی و پوشش گیاهی در پراکنندگی سالک نقش دارند. پزشکی و همکاران (۲۰۱۲) در چابهار، ارتباط بین وبا و دما، رطوبت بالا و بارش کم را بررسی کردند. اکبریور و همکاران (۲۰۱۲) با استفاده از مدل رگرسیون وزن‌دار جغرافیایی (GWR)، عوامل آب و هوایی مؤثر بر سالک در استان فارس را شناسایی کردند. کوردور و همکاران (۲۰۱۴) در جزیره ترینیداد نشان دادند که بارش، توپوگرافی و شاخص رطوبت بر شیوع لپتوسپیروز تأثیر دارند. محمدی نیا و همکاران (۲۰۱۵) نیز برتری مدل GWR را نسبت به رگرسیون حداقل مربعات معمولی (OLS) در بررسی لپتوسپیروز نشان دادند. آهنگرکانی و همکاران (۱۳۹۵) در شمال ایران از مدل‌های GWR و شبکه عصبی برای پیش‌بینی توزیع مکانی-زمانی لپتوسپیروز استفاده کردند. رجبی و همکاران (۲۰۱۴) با استفاده از شبکه عصبی مصنوعی، نقشه آسیب‌پذیری لیشمانیوز احشایی در آذربایجان شرقی را تهیه کردند. اخوان و همکاران (۱۳۹۲) نیز با استفاده از شبکه عصبی، میزان شیوع سالک در ایلام را پیش‌بینی کردند و نشان دادند که رطوبت و ارتفاع بیشترین اهمیت را دارند. کوچ (۲۰۱۶) با رویکردی چند سطحی، به تجزیه و تحلیل پیشگیری از سرطان پوست در آریزونا پرداخت. زیاری و همکاران (۲۰۱۵) نشان دادند که مناطق مرکزی ایران دارای خطر بیشتری برای سرطان پوست هستند. لورنزو و همکاران (۲۰۱۶) نیز الگوهای فضایی ملانوم در ایتالیا را بررسی کردند. کاتورانی و همکاران (۱۳۹۱) به بررسی عوامل اقلیمی و توپوگرافی مؤثر در سرطان پوست در کردستان پرداختند. کاظمی و همکاران (۱۳۹۴) فراوانی و توزیع جغرافیایی سرطان پوست در استان کردستان

را بررسی کردند. ویلر و همکاران (۲۰۱۳) ارتباط بین سرطان پوست و عوامل محیطی در انگلستان را بررسی کردند. افضلی و همکاران (۱۳۹۲) افزایش سرطان پوست در ایران را نشان دادند. مبارک‌آبادی و همکاران (۱۳۹۲) نشان دادند که سرطان پوست در مردان بیشتر از زنان است. زاریا و همکاران (۲۰۰۶) اثر تابش فرابنفش بر سرطان پوست را نشان دادند. پترا و همکاران (۲۰۰۸) تغییرات جغرافیایی سرطان پوست در استرالیا را بررسی کردند. شاسا و همکاران (۲۰۱۴) مناطق خطر سرطان پوست در فلوریدا را پیش‌بینی کردند. نورعلی‌نژاد (۱۳۹۳) نیز ویژگی‌های فردی بیماران سرطان پوست را بررسی کرد و نشان داد که تماس با نور آفتاب یکی از مهم‌ترین عوامل خطر است. مطالعات قبلی عمدتاً از مدل‌ها برای بیماری‌های واگیردار استفاده کرده‌اند و بر تعداد و نرخ سرطان پوست در مناطق مختلف تمرکز داشته‌اند. بیشتر تحقیقات بر عوامل فردی مانند جنس و سن متمرکز بوده‌اند و عوامل طبیعی بدون مدل‌سازی بررسی شده‌اند؛ بنابراین، تحقیق در مورد عوامل اقلیمی با استفاده از مدل‌سازی و اعتبارسنجی آن‌ها احساس می‌شود.

۳ مبانی نظری تحقیق

۳.۱ نرخ بروز بیماری

نرخ بروز بیماری به تعداد موارد جدید ابتلا در یک جمعیت مشخص طی یک بازه زمانی معین اشاره دارد. این شاخص با استفاده از رابطه روبرو محاسبه می‌شود:

$$I = n / N$$

که در آن n تعداد موارد جدید بیماری و N جمعیت در معرض خطر طی آن دوره زمانی است (Bonita et al., 2006).

۳.۲ سرطان پوست

سرطان‌ها به علت صدمه به DNA سلول‌ها ایجاد می‌شوند که می‌تواند ارثی یا ناشی از عوامل محیطی باشد. این صدمه موجب رشد غیر قابل کنترل سلول‌ها و تشکیل تومور می‌شود (رضایی، ۱۳۹۲). تومورها به دو نوع خوش‌خیم و بدخیم تقسیم می‌شوند. تومورهای بدخیم که به آن‌ها سرطان گفته می‌شود، می‌توانند به دیگر اعضا متاستاز کرده و منجر به مرگ شوند (پیروز، ۱۳۹۰). سرطان پوست به دو دسته اصلی تقسیم می‌شود: (۱) نوع غیرملانوم که شامل کارسینوم سلول سنگفرشی و کارسینوم سلول بازال است. این نوع شایع‌تر است؛ اما مرگ‌ومیر کمتری دارد و (۲) ملانوم که نرخ ابتلا به آن کمتر است؛ اما مرگ‌ومیر بیشتری دارد (العسیری، ۱۳۹۳).

عوامل ایجاد سرطان پوست

عوامل ایجاد سرطان پوست بر اساس مطالعات گذشته در سه دسته اصلی قرار می‌گیرند:

۱. عوامل طبیعی مانند تابش خورشید، دما، رطوبت، فشار، ارتفاع و تعداد روزهای یخبندان؛
۲. عوامل اجتماعی و شغلی شامل نژاد، پوشش، فرهنگ، سابقه خانوادگی و نوع شغل؛
۳. ویژگی‌های فردی مانند سن، جنس، بیماری‌های پوستی، سابقه پیوند، آگاهی فرد از بیماری و تغذیه

۳،۳ درونیابی و روش‌های آن

داده‌های جغرافیایی در GIS به دو دسته تقسیم می‌شوند: داده‌های گسسته که مرزهای آن‌ها به طور دقیق قابل تعریف هستند (مانند جاده‌ها) و داده‌های پیوسته که به طور نمونه برداشت می‌شوند و در مناطق غیر نمونه‌گیری قابل اندازه‌گیری نیستند. درونیابی به فرایند برآورد میزان متغیرهای پیوسته در مناطق غیر نمونه‌گیری شده گفته می‌شود (بلیانی و حکیم‌دوست، ۱۳۹۳). درونیابی داده‌های محیطی با استفاده از سه روش آمار کلاسیک، میان‌یابی قطعی و روش‌های زمین‌آمار قابل انجام است (خسروی و عباسی، ۱۳۹۵).

روش‌های زمین‌آمار به تحلیل داده‌های محیطی با توجه به ارتباط مکانی میان نمونه‌ها پرداخته و همبستگی مکانی آن‌ها را در نظر می‌گیرند. این تکنیک‌ها تحت عنوان نظریه متغیرهای ناحیه‌ای شناخته می‌شوند و به پیش‌بینی نقاط مجهول بر اساس خودهمبستگی میان نقاط اندازه‌گیری شده و ساختار فضایی آن‌ها می‌پردازند (خسروی و عباسی، ۱۳۹۵؛ بلیانی و حکیم‌دوست، ۱۳۹۳). یکی از مهم‌ترین روش‌های درونیابی، کریجینگ است که هدف آن یافتن وزن‌های آماری نمونه‌ها به گونه‌ای است که واریانس تخمین حداقل شود (بلیانی و حکیم‌دوست، ۱۳۹۳). انواع کریجینگ شامل کریجینگ معمولی، عام، ساده و انفصالی است (ترابی و همکاران، ۱۳۹۵).

۳،۴ تحلیل مکانی نرخ بروز بیماری سرطان پوست

در تحلیل‌های محیطی، داده‌ها معمولاً به دلیل موقعیت مکانی خود مستقل نیستند و دارای وابستگی فضایی هستند. تحلیل فضایی به بررسی الگوهای توزیع پدیده‌ها بر سطح زمین و قوانین فضایی حاکم بر آن‌ها می‌پردازد (بلیانی و حکیم‌دوست، ۱۳۹۳). تحلیل مکانی داده‌ها شامل سه رویکرد پردازشی، توصیفی و مدل‌سازی است که در مطالعات پیشین درباره سرطان پوست، بیشتر بر توصیف و تحلیل داده‌ها تمرکز شده است. برای پیش‌بینی و درک دقیق رفتار یک پدیده محیطی، نیاز به مدل‌سازی آن فرایند وجود دارد (خسروی و عباسی، ۱۳۹۵).

مدل‌سازی توزیع جغرافیایی بیماری‌ها با توجه به شرایط محیطی، ابزاری مؤثر برای برنامه‌ریزی، تخصیص منابع و مدیریت بهتر بیماری‌هاست (Burrough, 1986) گرچه تحلیل آماری داده‌های فضایی از دیرباز مورد توجه بوده، اما برای داده‌های حجیم مناسب نیست؛ در نتیجه، تحلیل مکانی به دلیل توانایی در کشف روابط، استخراج دانش و درک بهتر پدیده‌های جغرافیایی اهمیت بیشتری یافته است (Miller and Han, 2009) ویژگی‌هایی مانند هندسه و موقعیت مکانی داده‌ها، تحلیل آن‌ها را پیچیده می‌کند؛ از این رو، استفاده از روابط و الگوریتم‌های مکانی ضروری است. از روش‌های رایج مدل‌سازی مکانی می‌توان به رگرسیون وزندار جغرافیایی و شبکه‌های عصبی مصنوعی اشاره کرد.

۳,۴,۱ رگرسیون وزندار جغرافیایی

رگرسیون وزندار جغرافیایی (GWR) روشی مناسب برای مدل‌سازی روابط مکانی است که در آن فرض بر یکنواختی روابط بین متغیرها در کل منطقه وجود ندارد. این روش با در نظر گرفتن ناهمگنی مکانی، امکان برآورد جداگانه مدل برای هر منطقه را فراهم می‌سازد (بلیانی و حکیم‌دوست، ۱۳۹۳؛ عسگری، ۱۳۹۰). در GWR، برای هر نقطه، مدل بر اساس داده‌های اطراف آن برآورد می‌شود و داده‌های نزدیک‌تر وزن بیشتری دریافت می‌کنند. دو نوع هسته در GWR وجود دارد: هسته ثابت که پهنای باند مشخصی دارد و برای همه نقاط یکسان است. هسته تطبیقی که پهنای باند آن بسته به تراکم داده‌ها تغییر می‌کند، به طوری که در مناطق پراکنده پهنای باند بیشتر و در مناطق متراکم کمتر می‌شود (بلیانی و حکیم‌دوست، ۱۳۹۳). این روش به کمک ماتریس وزن‌ها، وابسته به موقعیت مکانی نقاط، به برآورد پارامترهای مدل می‌پردازد و از این طریق می‌توان روابط منطقه‌ای دقیق‌تری بین متغیرهای مستقل و وابسته برقرار کرد.

۳,۴,۲ شبکه عصبی مصنوعی پرسپترون چندلایه

شبکه عصبی مصنوعی پرسپترون چندلایه (MLP) یکی از رایج‌ترین مدل‌های شبکه عصبی است که از ساختار مغز انسان الهام گرفته شده و به‌ویژه در مدل‌سازی مسائل پیچیده کاربرد دارد (Hornik, ۱۹۸۹). این نوع شبکه از لایه‌های پی‌درپی نورون‌ها تشکیل شده است که شامل یک لایه ورودی، یک یا چند لایه پنهان و یک لایه خروجی است. همه نورون‌های هر لایه به‌طور کامل به نورون‌های لایه بعد متصل‌اند و مسیر عبور سیگنال‌ها همواره از ورودی به خروجی است؛ بنابراین، بازخوردی در ساختار وجود ندارد و شبکه به صورت پیش‌رو عمل می‌کند. در این شبکه‌ها دو نوع سیگنال وجود دارد: سیگنال‌های تابعی که در مسیر رفت از ورودی به خروجی جریان دارند و بر اساس ورودی و پارامترهای نورون محاسبه می‌شوند. سیگنال‌های خطایی که در مسیر برگشت از خروجی به ورودی حرکت می‌کنند و برای تنظیم وزن‌ها در فرایند آموزش به کار می‌روند (منهاج، ۱۳۹۵). این ساختار باعث می‌شود پرسپترون چندلایه برای شناسایی الگوها و روابط پیچیده بین داده‌ها بسیار مؤثر باشد.

۴ روش تحقیق

۴,۱ منطقه مورد مطالعه

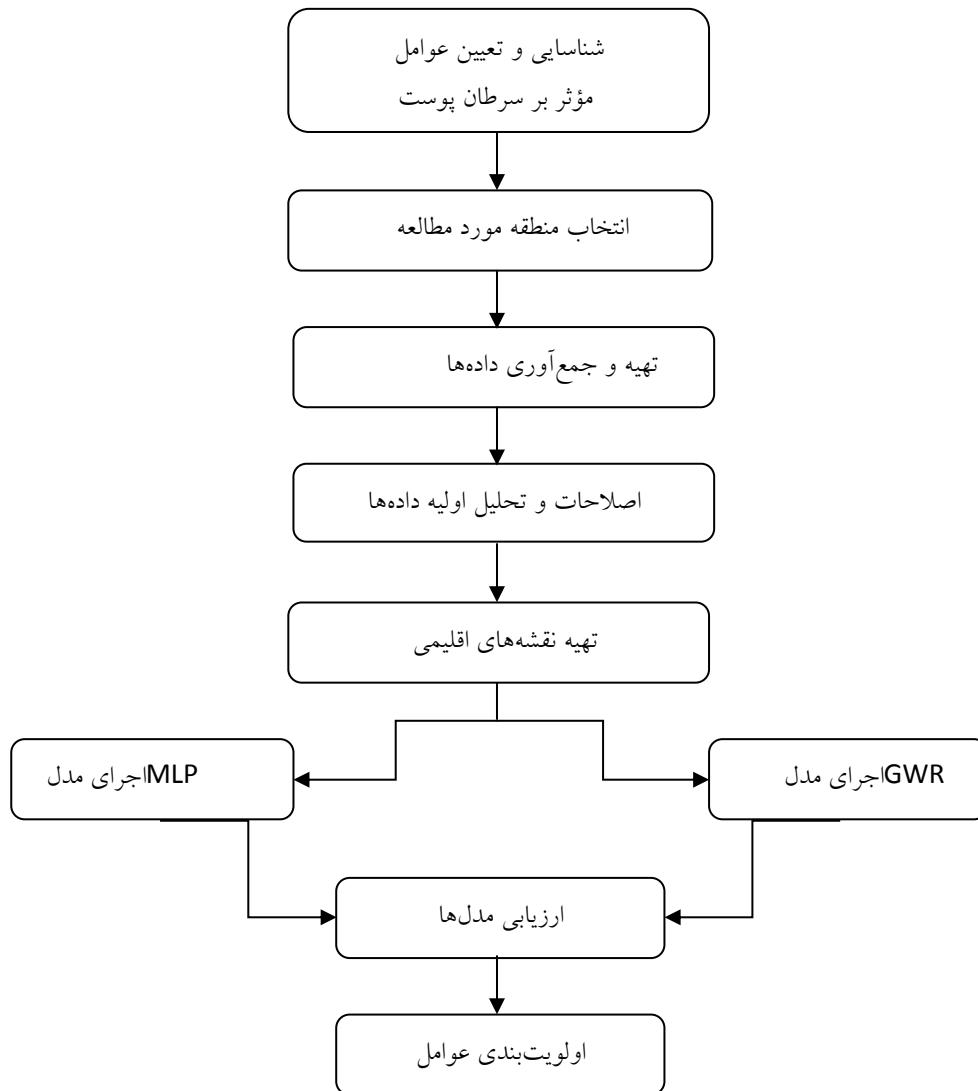


کشور ایران با وسعتی بیش از ۱.۶ میلیون کیلومتر مربع، در نیمکره شمالی و منطقه معتدل قرار دارد و به دلیل ویژگی‌های جغرافیایی مانند کوهستان‌ها، بیابان‌ها، مجاورت با دریاها و نزدیکی به مناطق اقلیمی متنوع، دارای اقلیم‌های گوناگونی است. این پژوهش بر اساس تقسیمات کشوری ایران در سطح شهرستان‌های کل کشور انجام شده است.

شکل (۱) نقشه موقعیت منطقه مورد مطالعه

۴,۲ روش انجام تحقیق

در این تحقیق، ابتدا با مرور مطالعات پیشین، عوامل محیطی و اقلیمی مؤثر بر سرطان پوست شناسایی شدند و شهرستان‌های کشور به‌عنوان منطقه مطالعه انتخاب شدند. داده‌های مربوط به بیماران از اداره سرطان مرکز بیماری‌های غیرواگیر وزارت بهداشت و داده‌های اقلیمی از سازمان هواشناسی کل کشور گردآوری شد. در ادامه، داده‌ها اصلاح، تجزیه و تحلیل شده و نرخ بروز سرطان پوست برای هر شهرستان محاسبه گردید. داده‌های اقلیمی ماهانه نیز به میانگین سالانه تبدیل و به صورت لایه‌های اطلاعات مکانی در قالب GIS سازمان‌دهی شدند. سپس با استفاده از روش درون‌یابی، نقشه‌های اقلیمی تهیه و از طریق روش Zonal Statistics، میانگین پارامترهای اقلیمی برای هر شهرستان استخراج شد. پس از فازی‌سازی داده‌ها، دو مدل شبکه عصبی مصنوعی پرسپترون چندلایه و رگرسیون وزندار جغرافیایی برای مدل‌سازی نرخ بروز سرطان پوست به کار گرفته شدند. در پایان، پارامترهای اقلیمی مؤثر بر بیماری بر اساس میزان تأثیر و نوع ارتباط با نرخ بروز سرطان پوست اولویت‌بندی گردیدند. مراحل روش تحقیق در قالب نمودار در شکل (۲) نمایش داده شده است:



شکل (۲) روش انجام تحقیق (نگارنده)

۴,۳ داده‌های استفاده شده در تحقیق

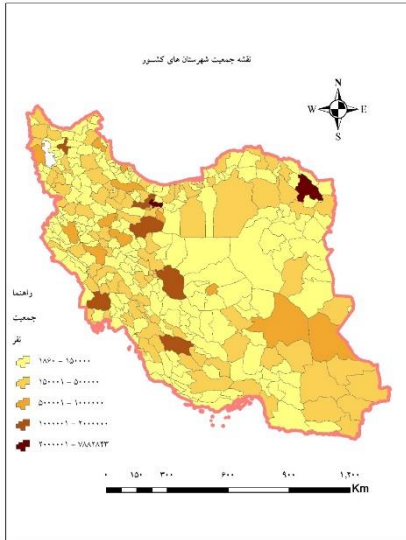
بر اساس تحقیقات انجام شده، عوامل طبیعی، محیط اجتماعی و کاری و ویژگی‌های فردی در ایجاد سرطان پوست مؤثر هستند؛ بنابراین داده‌های مورد استفاده شامل داده‌های تقسیمات کشوری، جمعیتی، بیماری سرطان پوست و عوامل طبیعی مؤثر بر ایجاد این بیماری می‌باشند.

۴,۳,۱ داده‌های تقسیمات کشوری

داده‌های تقسیمات کشوری از دفتر تقسیمات کشوری وزارت کشور و مرکز آمار ایران تهیه شدند. این داده‌ها به فرمت‌های Shp (از وزارت کشور) و Excel (از مرکز آمار ایران) موجود بودند. از فرمت Excel برای اصلاح آدرس‌دهی داده‌های بیماری استفاده شد.

۴,۳,۲ داده‌های جمعیت

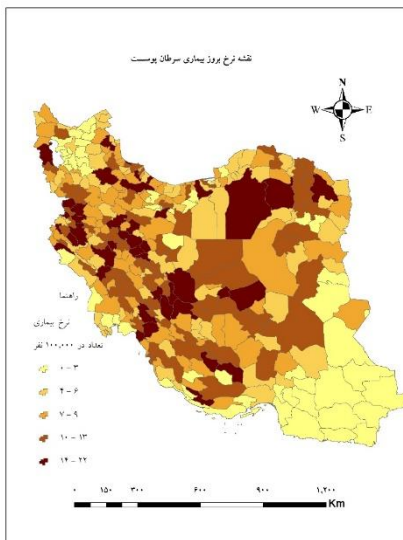
داده‌های جمعیت بر اساس سرشماری عمومی نفوس و مسکن سال ۱۳۹۵ از مرکز آمار ایران تهیه شدند. این داده‌ها در فرمت Excel به Shp شهرستان‌ها پیوند داده شدند و برای محاسبه نرخ بروز بیماری استفاده گردید.



شکل (۳) نقشه جمعیت

۴,۳,۳ داده‌های بیماری

داده‌های بیماری سرطان پوست از اداره سرطان وزارت بهداشت برای تهیه شدند. این داده‌ها شامل اطلاعاتی مانند استان، شهر، و آدرس بیمار بودند. مشکلات داده‌ها شامل فیلدهای خالی و اشتباهات آدرسی بود که پس از اصلاحات، آدرس‌ها در سطح استان و شهرستان اصلاح شدند. برای حذف تأثیر جمعیت در مقایسه شهرستان‌ها، نرخ بروز بیماری به ازای هر 100,000 نفر محاسبه شد، بدین صورت که تعداد بیماران هر شهرستان تقسیم بر جمعیت آن شهرستان و ضرب در ۱۰۰,۰۰۰ شد (Bonita et al., 2006).

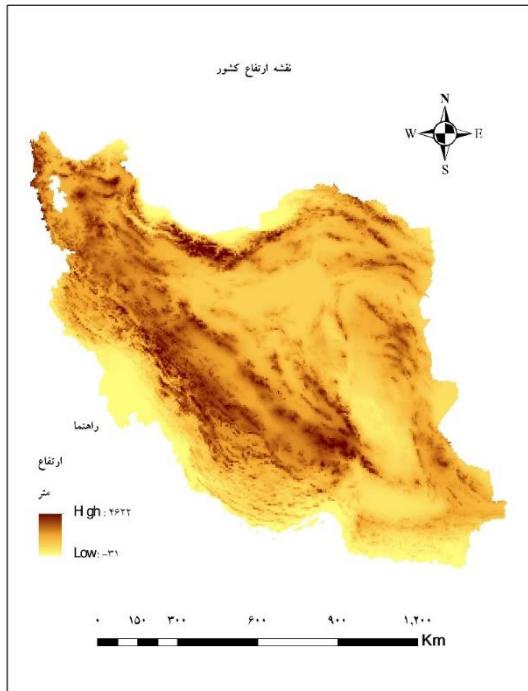


شکل (۴) نقشه نرخ بروز سرطان پوست

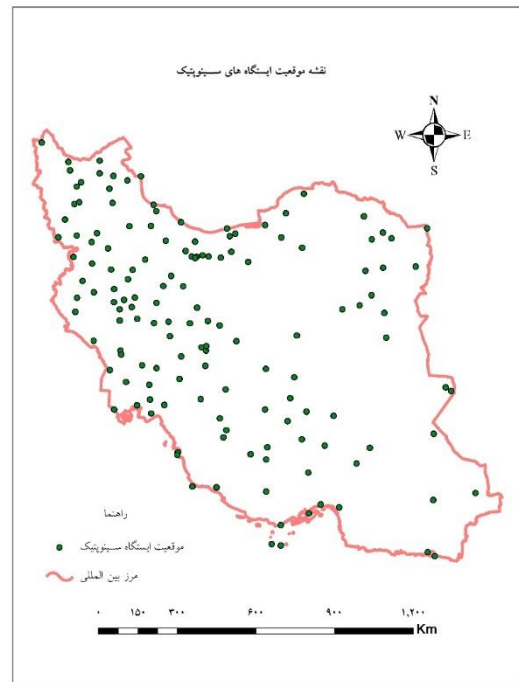
۴,۳,۴ داده‌های اقلیمی و محیطی

داده‌های اقلیمی مؤثر بر سرطان پوست شامل ارتفاع متوسط، ساعات آفتابی، بارش، روزهای برف و تگرگ، و روزهای گردوغبار از سالنامه‌های آماری هواشناسی کشور و ایستگاه‌های سینوپتیک استخراج شدند. این داده‌ها به صورت ماهانه و در فرمت txt برای ۱۴۶ ایستگاه موجود بودند. برای تبدیل داده‌ها به سطح پیوسته، از روش‌های درونیابی کریجینگ استفاده شد تا نقشه‌های پیوسته برای کل کشور تهیه گردد.

یکی از مهم‌ترین روش‌ها برای ارزیابی دقت مدل‌های میان‌یابی، ریشه میانگین مربع خطا (RMSE) است که هر چه مقدار آن کمتر باشد، دقت مدل بالاتر است (بلیانی و حکیم‌دوست، ۱۳۹۳). RMSE میزان خطا بین مقادیر پیش‌بینی شده و مقادیر اندازه‌گیری شده را اندازه‌گیری می‌کند. اعتبارسنجی روش‌ها بر اساس RMSE انجام شده و روش‌های کریجینگ با کمترین مقدار RMSE به‌عنوان روش نهایی انتخاب شده‌اند. برای تهیه نقشه متوسط ارتفاع، از داده‌های DEM تصاویر ماهواره‌ای SRTM با تفکیک‌پذیری ۹۰ متر استفاده شده است.



شکل (۶) نقشه ارتفاع کشور



شکل (۵) نقشه موقعیت ایستگاه‌های سینوپتیک کشور

۴,۴ مدل‌ها

فرایند آماده‌سازی داده‌ها یکی از مراحل حیاتی در اجرای مدل‌های تحلیلی به شمار می‌رود و شامل مراحل مختلفی مانند تحلیل آماری مکانی (زونی)، بررسی هم‌خطی چندگانه میان متغیرها و نرمال‌سازی داده‌هاست که به منظور اطمینان از دقت و صحت تحلیل‌های بعدی انجام می‌شود.

در گام نخست، داده‌های مربوط به نرخ بروز سرطان پوست به‌صورت عددی برای هر شهرستان جمع‌آوری شده‌اند و به‌منظور تحلیل مکانی عوامل مؤثر، نقشه‌های مربوط به متغیرهای طبیعی با روش درون‌یابی تهیه شدند. برای اختصاص مقادیر در سطح شهرستان، از روش Zonal Statistics استفاده شد. این روش با اعمال توابع آماری نظیر میانگین و میانه بر روی لایه‌های مکانی، مقادیر متغیرهای اقلیمی و ارتفاع را به تفکیک هر محدوده شهرستان محاسبه می‌کند (نوحه‌گر و اکبری‌نسب، ۱۳۹۲).

در ادامه، جهت استفاده از مدل رگرسیون وزن‌دار جغرافیایی (GWR)، بررسی هم‌خطی چندگانه میان متغیرهای مستقل ضروری بود. هم‌خطی زمانی ایجاد می‌شود که دو یا چند متغیر مستقل دارای همبستگی بالایی با یکدیگر باشند که این موضوع باعث اختلال در نتایج مدل رگرسیون می‌شود. برای شناسایی این مشکل، از شاخص فاکتور تورم واریانس (Variance Inflation Factor - VIF) استفاده شد. اگر مقدار VIF بزرگ‌تر از ۷.۵ باشد، بیانگر هم‌خطی شدید بین متغیرها است که مانع اجرای صحیح مدل می‌شود (جعفری و حبیبی ماچیان، ۱۳۹۳؛ عسگری، ۱۳۹۰). در نهایت، برای اجرای مدل پرسپترون چندلایه (MLP) در محیط نرم‌افزاری MATLAB، نیاز به نرم‌ساز داده‌ها وجود دارد. نرم‌ساز داده‌ها باعث می‌شود تمامی ورودی‌های مدل دارای مقیاس یکسان باشند و دقت آموزش مدل افزایش یابد. این فرایند در محیط نرم‌افزاری ArcGIS با استفاده از روش عضویت فازی خطی انجام گرفت. در این روش، حداکثر مقدار داده‌ها دارای عضویت فازی ۱ و حداقل مقدار دارای عضویت فازی صفر هستند و سایر مقادیر بین این دو مقدار قرار می‌گیرند. در این تحقیق، کلیه لایه‌ها اعم از نرخ بیماری و عوامل طبیعی مؤثر بر سرطان پوست، به‌صورت فازی‌سازی شده آماده شدند.

۴.۴.۱ مدل‌سازی بیماری سرطان پوست

این تحقیق با هدف بررسی رابطه میان عوامل طبیعی و اقلیمی با نرخ بروز بیماری سرطان پوست در سطح شهرستان‌های ایران انجام شده است. برای دستیابی به این هدف، پس از تهیه نقشه‌های اطلاعاتی از متغیرهای اقلیمی و محیطی، دو مدل تحلیلی شامل رگرسیون وزن‌دار جغرافیایی (GWR) و شبکه عصبی پرسپترون چندلایه (MLP) اجرا و نتایج حاصل از آن‌ها تحلیل و ارزیابی گردید. در ابتدا، به‌منظور تحلیل مکانی و درک توزیع فضایی تأثیر عوامل طبیعی بر نرخ بروز سرطان پوست، از مدل رگرسیون وزن‌دار جغرافیایی (GWR) استفاده شد. در این مدل، نرخ بروز بیماری به‌عنوان متغیر وابسته و متغیرهای طبیعی شامل ساعات آفتابی، میزان بارش، تعداد روزهای برفی، گردوغبار، و ارتفاع به‌عنوان متغیرهای مستقل در نظر گرفته شدند. داده‌های مربوطه پس از آماده‌سازی در نرم‌افزارهای Excel، SPSS و ArcGIS به مدل GWR معرفی گردیدند. همچنین برای انتخاب مدل هسته مکانی، از هسته تطبیقی استفاده شد تا دقت مدل در تحلیل‌های مکانی افزایش یابد. خروجی مدل، نقشه‌هایی بود که ضرایب تأثیرگذاری هر عامل را به‌صورت مکانی نشان می‌داد و مناطق با ریسک بالاتر را مشخص می‌کرد. در مرحله بعد، به‌منظور مدل‌سازی پیش‌بینی‌پذیرتر و غیرخطی‌تر روابط میان متغیرها، از شبکه عصبی پرسپترون چندلایه (MLP) استفاده شد. داده‌ها پس از فازی‌سازی در محیط ArcGIS، وارد نرم‌افزار MATLAB گردیدند. در این مدل نیز متغیرهای طبیعی شامل مجموع ساعات آفتابی، مجموع بارش، تعداد روزهای برف و تگرگ، تعداد روزهای گردوغبار و متوسط ارتفاع به‌عنوان ورودی و نرخ بروز سرطان پوست به‌عنوان خروجی به مدل معرفی شدند. فرایند آموزش مدل با استفاده از تقسیم‌بندی داده‌ها به دو دسته انجام شد: ۸۰ درصد داده‌ها برای آموزش و ۲۰ درصد باقی‌مانده برای آزمون و اعتبارسنجی مدل اختصاص یافت. مدل نهایی توانست با دقت مطلوبی نرخ بروز سرطان پوست را در شهرستان‌های مختلف کشور مدل‌سازی کرده و الگوهای پنهان در داده‌ها را استخراج کند. ترکیب این دو روش، امکان مقایسه‌ی قدرت تبیین و پیش‌بینی آن‌ها را

فراهم آورد و نقش هریک از عوامل اقلیمی را در شکل‌گیری سرطان پوست با دقت بیشتری نمایان ساخت. این تحقیق با هدف بررسی ارتباط عوامل طبیعی مؤثر بر بیماری سرطان پوست و نرخ بروز آن انجام شد. برای این منظور، نقشه‌های اطلاعاتی عوامل مختلف در سطح شهرستان‌ها تهیه و دو مدل رگرسیون وزندار جغرافیایی و شبکه عصبی پرسپترون چندلایه اجرا و نتایج آن‌ها تحلیل و ارزیابی گردید.

۴,۴,۲ اعتبار سنجی مدل‌ها

برای بررسی کارایی مدل‌ها و انتخاب مدل بهتر، از ضریب همبستگی R استفاده شد. این ضریب ارتباط بین داده‌های نرخ سرطان پوست و مقادیر پیش‌بینی شده مدل‌ها را نشان می‌دهد. هرچه این ضریب به عدد یک نزدیک‌تر باشد، دقت مدل بیشتر است (جعفری و حبیبی ماچانی، ۱۳۹۳). تفسیر ضریب همبستگی به شرح زیر است:

۱۰٪ تا ۲۹٪ = همبستگی پایین، ۳۰٪ تا ۴۹٪ = همبستگی متوسط، ۵۰٪ تا ۱ = همبستگی بالا

۴,۴,۳ اولویت‌بندی عوامل ایجاد سرطان پوست

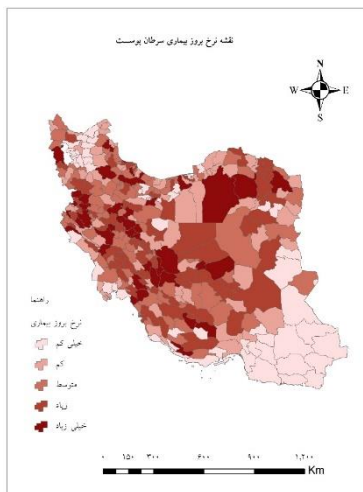
برای تعیین اهمیت عوامل مؤثر در مدل‌سازی سرطان پوست، از آنالیز حساسیت استفاده شد. در این روش، با حذف هر عامل از مدل و محاسبه مقدار R ، اهمیت هر عامل مشخص گردید. عاملی که بیشترین کاهش در مقدار R را ایجاد کند، به‌عنوان مهم‌ترین عامل شناخته شد.

۵ پیاده‌سازی و نتایج

در این بخش، نتایج حاصل از آماده‌سازی و تحلیل داده‌ها و اجرای مدل‌های تحقیق ارائه می‌شود. پس از شناسایی و اصلاح عوامل مؤثر بر سرطان پوست، داده‌ها برای مدل‌سازی آماده شدند و نتایج اجرای مدل‌ها نمایش داده شد. همچنین برای اعتبارسنجی مدل‌ها از ضریب همبستگی R استفاده شده و در پایان، متغیرهای مؤثر بر سرطان پوست اولویت‌بندی گردیده‌اند.

۵,۱ تحلیل داده‌های بیماری سرطان پوست

پس از اصلاح آدرس‌های ناقص حدود ۲۴,۵٪ به دلیل نداشتن آدرس حذف شدند. سپس با پیوند اطلاعات بیماران به داده‌های مکانی شهرستان‌ها، نرخ بروز بیماری برای هر شهرستان محاسبه و نقشه‌ای از آن تهیه شد. این داده‌ها برای استفاده در مدل شبکه عصبی پرسپترون چندلایه، فازی‌سازی شدند.



شکل (۷) نقشه نرخ بروز بیماری سرطان پوست (فازی شده)

۵،۲ تحلیل داده‌های اقلیمی و طبیعی

جدول (۱) نتایج مقادیر RMSE

مقدار RMSE	روش کریجینگ	متغیر اقلیمی
۰.۶۳	معمولی	تعداد روز برف و تگرگ
۰.۷۰	ساده	
۰.۶۳	عام	
۰.۷۲	انفصالی	
۲.۲۲	معمولی	تعداد روز گرد و غبار
۲.۱۷	ساده	
۲.۲۲	عام	
۲.۱۵	انفصالی	
۱۴.۷۶	معمولی	مجموع ساعات آفتابی
۱۵.۵۴	ساده	
۱۷.۱۹	عام	
۱۶.۳۹	انفصالی	
۱۲.۰۹	معمولی	مجموع بارش
۱۵.۴۴	ساده	
۱۰.۷۷	عام	
۱۷.۸۷	انفصالی	

برای تحلیل داده‌های اقلیمی، میانگین سالانه پارامترهای اقلیمی برای ایستگاه‌های سینوپتیک محاسبه و به لایه‌های نقطه‌ای با اطلاعات مکانی تبدیل شد. جهت ورود این داده‌ها به مدل‌ها، با استفاده از روش‌های درونیابی زمین‌آماری کریجینگ، نقشه‌های پیوسته اقلیمی برای کل کشور تهیه گردید. انواع روش‌های کریجینگ (معمولی، عام، ساده و انفصالی) به کار گرفته شده و با معیار RMSE ارزیابی شدند. روشی که کمترین RMSE را داشت، به‌عنوان روش نهایی انتخاب شد. در جدول (۱) نتایج مقادیر RMSE برای انواع روش‌های کریجینگ آورده شده است.

برای درونیابی پارامترهای اقلیمی، روش‌های مختلف کریجینگ بر اساس مقدار RMSE ارزیابی شدند. برای تعداد روز برف و تگرگ، دو روش کریجینگ معمولی و عام با RMSE برابر ۰،۶۳ بهترین عملکرد را داشتند که در نهایت روش معمولی انتخاب شد. در مورد گردوغبار، کریجینگ انفصالی با RMSE معادل ۲،۱۵ دقیق‌ترین روش

بود. برای مجموع ساعات آفتابی، کریجینگ معمولی با RMSE برابر ۱۴،۶۷ و برای مجموع بارش، کریجینگ عام با RMSE برابر ۱۰،۷۷ بهترین دقت را نشان دادند و به‌عنوان روش‌های نهایی درونیابی انتخاب شدند.

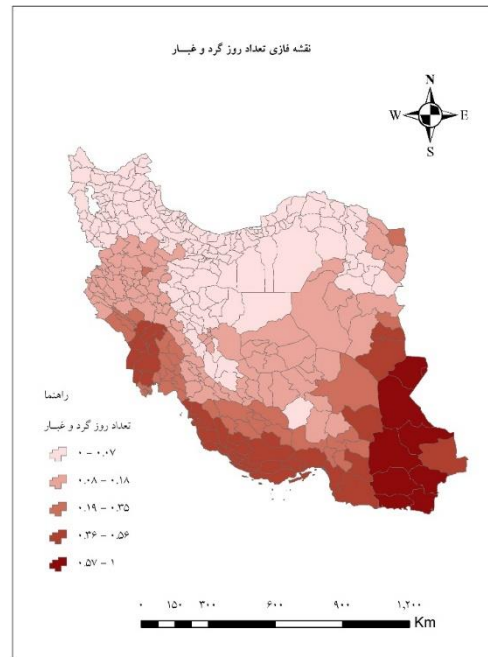
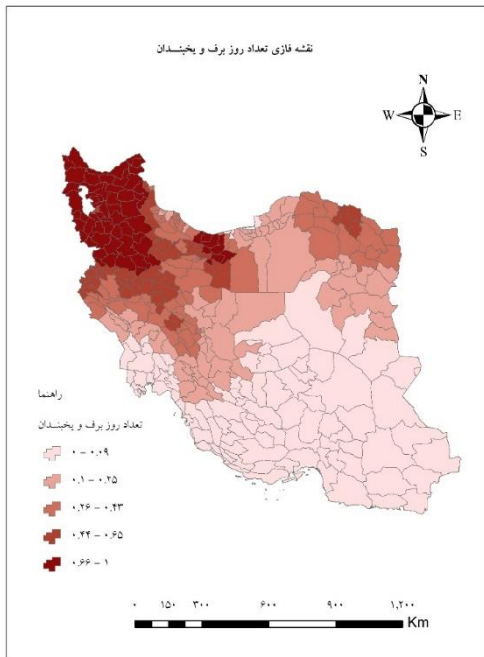
از آنجا که نقشه‌های اقلیمی برای کل کشور تهیه شده‌اند و تحلیل تحقیق در سطح شهرستان است، برای اختصاص مقدار مشخصی از هر متغیر اقلیمی به هر شهرستان، از تحلیل آماری زونی استفاده شد.

جدول (۲) مقدار VIF از متغیرهای مورد استفاده

مقدار VIF	متغیر مستقل	متغیر وابسته
۱.۹۴	تعداد روز برف و تگرگ	نرخ شیوع بیماری سرطان پوست
۲.۳۱	تعداد روز گرد و غبار	
۳.۶۸	مجموع ساعات آفتابی	
۲.۲۸	مجموع بارش	
۲.۲۸	میانگین ارتفاع	

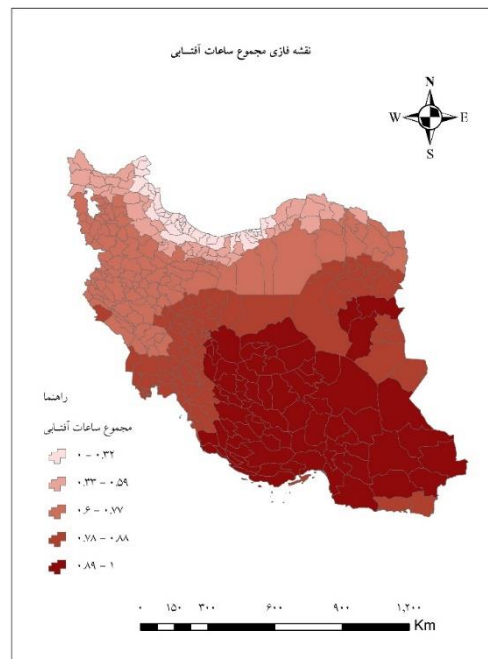
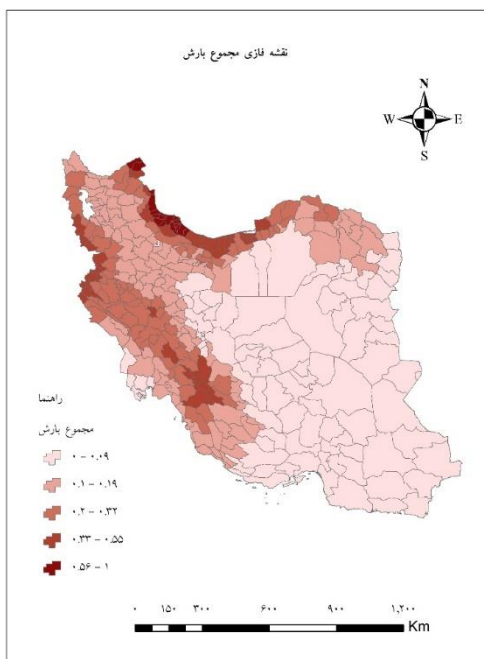
پیش از فازی‌سازی داده‌های طبیعی، شاخص VIF برای بررسی چند هم‌خطی بین متغیرهای مستقل محاسبه شد. با توجه به اینکه همه مقادیر VIF کمتر از ۷.۵ بودند، مستقل بودن متغیرهای اقلیمی تأیید شد. نتایج این بررسی در جدول (۲) نشان داده شده‌اند.

در نهایت، نقشه‌های زونی، فازی‌سازی شده و نقشه‌های اقلیمی نهایی برای ورود به مدل آماده شدند. در ادامه نقشه‌های اقلیمی نهایی ورودی به مدل نشان داده شده‌اند:

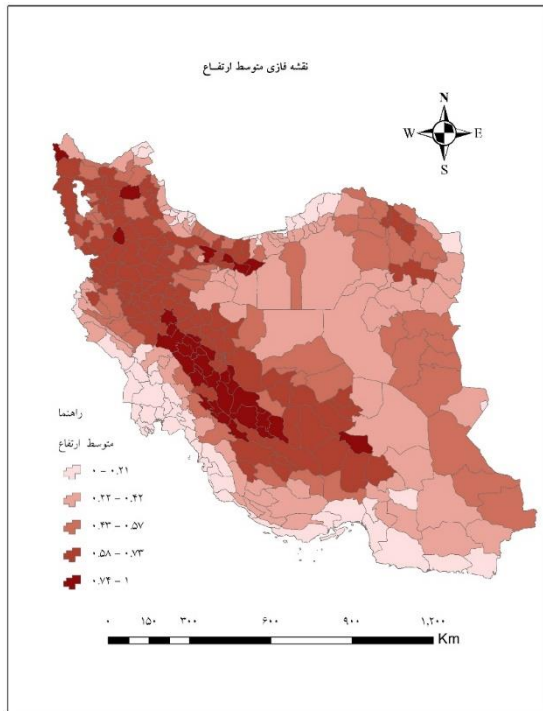


شکل (۹) نقشه فازی متغیر اقلیمی تعداد روز برف و یخبندان
شکل (۱۰) نقشه فازی متغیر اقلیمی مجموع ساعات آفتابی

شکل (۸) نقشه فازی متغیر اقلیمی تعداد روز گردوغبار



شکل (۱۱) نقشه فازی متغیر اقلیمی مجموع بارش



نقشه ارتفاعی بر اساس مدل رقومی ارتفاع (DEM) تهیه شد. سپس با استفاده از تحلیل آماری زونی، میانگین ارتفاع هر شهرستان محاسبه و داده‌ها فازی‌سازی شدند. نقشه نهایی ارتفاع برای ورود به مدل آماده گردید.

شکل (۱۲) نقشه فازی متغیر میانگین ارتفاع

۵,۳ اجرای مدل‌ها

پس از انجام عملیات زونی، بررسی هم‌خطی و فازی‌سازی داده‌ها، شرایط لازم برای اجرای مدل‌سازی نرخ سرطان پوست فراهم شد. نتایج مدل‌های رگرسیون و زنده‌دار جغرافیایی و شبکه عصبی پرسپترون چندلایه در ادامه ارائه شده است.

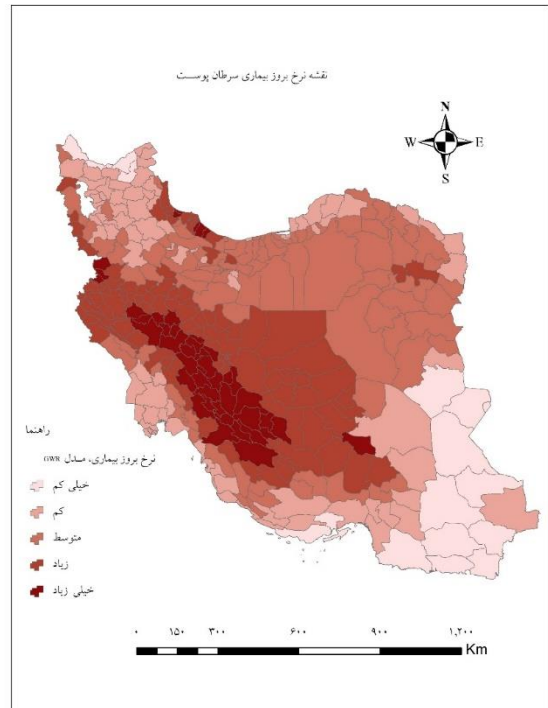
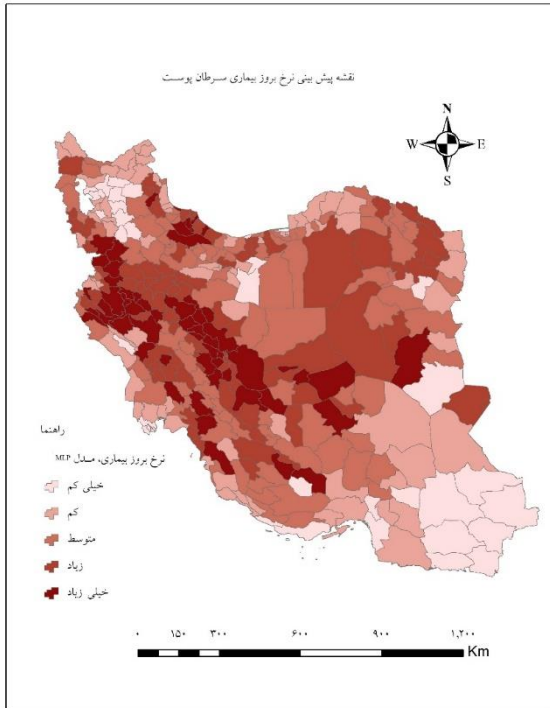
۵,۳,۱ رگرسیون و زنده‌دار جغرافیایی

مدل رگرسیون و زنده‌دار جغرافیایی با استفاده از هسته تطبیقی برای بررسی رابطه بین عوامل طبیعی و نرخ بروز سرطان پوست اجرا شد. نتایج نشان داد که نرخ بروز بیماری بیشتر به صورت منطقه‌ای است و مناطق هم‌جوار نرخ‌های مشابهی دارند. بیشترین نرخ‌ها در شهرستان‌های جنوب غرب، شمال و غرب کشور دیده می‌شود که این مناطق ارتفاع و بارش بالایی دارند.

۵,۳,۲ شبکه عصبی پرسپترون چندلایه

در مدل‌سازی با شبکه عصبی پرسپترون چندلایه، ۸۰٪ داده‌ها برای آموزش، ۱۰٪ برای ارزیابی و ۱۰٪ برای آزمون مدل استفاده شد. برخلاف رگرسیون و زنده‌دار جغرافیایی، نتایج این مدل کمتر حالت منطقه‌ای دارند. نرخ‌های بالای

سرطان پوست علاوه بر جنوب غرب، شمال و غرب، در نواحی مرکزی کشور نیز مشاهده می شود. در این مناطق، بارش در شمال، غرب و جنوب غرب زیاد و ساعات آفتابی در مرکز، غرب و جنوب غرب بالا است.



شکل (۱۴) نقشه مدل سازی نرخ بروز بیماری سرطان پوست با استفاده از مدل شبکه عصبی پرسپترون چندلایه

شکل (۱۳) نقشه مدل سازی نرخ بروز بیماری سرطان پوست با استفاده از مدل رگرسیون وزندار جغرافیایی

۵,۴ اعتبارسنجی مدل ها

برای ارزیابی دقت مدل سازی سرطان پوست، هر دو مدل با استفاده از ضریب همبستگی R مورد اعتبارسنجی قرار گرفتند. نتایج این ارزیابی در جدول (۳) ارائه شده است.

جدول (۳) مقدار R هر دو مدل رگرسیون وزندار جغرافیایی و شبکه عصبی پرسپترون چندلایه

مقدار R	مدل
۰.۴۹	رگرسیون وزندار جغرافیایی
۰.۷۴	شبکه عصبی پرسپترون چندلایه

نتایج ارزیابی نشان می دهد که مدل شبکه عصبی پرسپترون چندلایه دقت بالاتری نسبت به رگرسیون وزندار جغرافیایی در مدل سازی نرخ بروز سرطان پوست در سطح شهرستان ها دارد. در حالی که رگرسیون وزندار در مرز همبستگی متوسط و بالا قرار دارد، شبکه عصبی از همبستگی بالا برخوردار بوده و عملکرد دقیق تری ارائه می دهد.

۵.۵ اولویت بندی عوامل ایجاد سرطان پوست

برای تعیین میزان اهمیت عوامل مؤثر بر سرطان پوست، از آنالیز حساسیت استفاده شد. در این روش، با حذف جداگانه هر عامل و محاسبه ضریب همبستگی R، تأثیر آن بر دقت مدل بررسی شد. عاملی که بیشترین کاهش در R را ایجاد کرد، مهم ترین متغیر شناخته شد. این تحلیل برای هر دو مدل GWR و MLP انجام شد و نتایج اولویت بندی عوامل در جدول (۴-۴) ارائه گردیده است.

جدول (۴) مقدار نتایج آنالیز حساسیت هر دو مدل GWR و MLP

مقدار R		متغیر مستقل حذف شده از مدل
مدل GWR	مدل MLP	
۰.۴	۰.۶	مجموع بارش
۰.۴۳	۰.۶۵	مجموع ساعات آفتاب
۰.۴۵	۰.۶۸	میانگین ارتفاع
۰.۴۷	۰.۷۱	تعداد روز گردوغبار
۰.۴۸	۰.۷۳	تعداد روز برف و تگرگ

بر اساس جدول (۴)، مهم ترین متغیرهای مؤثر بر سرطان پوست به ترتیب اولویت عبارت اند از: مجموع بارش، مجموع ساعات آفتابی، میانگین ارتفاع، تعداد روزهای گردوغبار و تعداد روزهای برف و تگرگ.

۶ بحث و نتیجه گیری

با توجه به شیوع بالای سرطان پوست و هزینه های سنگین آن، تحقیق در این زمینه ارزشمند اما دشوار است، به ویژه به دلیل تأثیر عوامل طبیعی و دشواری در جمع آوری داده ها. برخلاف مطالعات قبلی که بر ویژگی های فردی تمرکز داشتند، این تحقیق عوامل جغرافیایی را بررسی کرده و با دو روش مدلسازی نرخ سرطان پوست انجام شده است. نتایج به دست آمده می تواند برای پیشگیری، هم برای مردم در معرض خطر و هم برای مسئولان سلامت مفید باشد.

۶.۱ بررسی فرضیات

با استناد به نتایج به دست آمده، امکان پاسخ گویی به سؤالات تحقیق و بررسی صحت فرضیات فراهم شده است.

فرضیه اول:

فرضیه اول تأیید شد؛ با استفاده از مدل های GWR و MLP، می توان نرخ بروز سرطان پوست را بر اساس متغیرهای اقلیمی مدلسازی کرد و نتایج ارزیابی نشان می دهد که هر دو مدل قادر به انجام مدلسازی با دقت قابل قبول هستند.

فرضیه دوم:

فرضیه دوم تأیید شد؛ مدل شبکه عصبی پرسپترون چندلایه (MLP) با دقت بالاتر ($R=0.74$) نسبت به مدل رگرسیون وزندار جغرافیایی ($R=0.49$) نرخ بروز سرطان پوست را بهتر مدلسازی می کند.

فرضیه سوم:

فرضیه سوم تأیید شد؛ با توجه به مدل‌سازی نرخ ابتلا به سرطان پوست با استفاده از مدل‌های MLP و GWR بر اساس عوامل طبیعی، می‌توان نتیجه گرفت که عوامل محیطی و مکانی بر ایجاد این بیماری مؤثر هستند.

فرضیه چهارم:

فرضیه چهارم رد شد؛ بر اساس جدول ۴-۴، عامل اقلیمی ساعات آفتابی دومین اولویت را دارد و نه بیشترین تأثیر در نرخ بروز سرطان پوست.

۶,۲ نتیجه‌گیری

هدف این تحقیق شناسایی و مدل‌سازی عوامل طبیعی مؤثر بر بروز سرطان پوست بود. از دو مدل شبکه عصبی پرسپترون چندلایه (MLP) و رگرسیون وزندار جغرافیایی (GWR) برای این منظور استفاده شد. نتایج نشان داد مدل MLP با دقت بالاتر ($R=0.74$) عملکرد بهتری نسبت به مدل GWR دارد.

آنالیز حساسیت نشان داد که عوامل مؤثر بر سرطان پوست به ترتیب اهمیت عبارت‌اند از: مجموع بارش، مجموع ساعات آفتابی، میانگین ارتفاع، تعداد روز گردوغبار و تعداد روز برف و تگرگ.

محدودیت‌های این تحقیق شامل مشکلات در جمع‌آوری داده‌های سرطان پوست به دلیل ناقص بودن گزارش‌ها و امکانات بهداشتی ناکافی، به ویژه در مناطق مرزی مانند سیستان و بلوچستان است. بسیاری از بیماری‌ها در این مناطق ثبت نمی‌شوند یا به اشتباه به استان‌های دیگر گزارش می‌شوند. دیگر محدودیت‌ها شامل قوانین سخت برای دسترسی به داده‌ها و تغییرات در تقسیمات سیاسی و اداری کشور است که باعث ناهماهنگی در داده‌ها می‌شود. همچنین، آمارهای جمعیتی به‌طور معمول تنها هر ده سال یک‌بار به‌روزرسانی می‌شوند.

۶,۳ پیشنهادها

پیشنهادها تحقیق شامل موارد زیر است:

ایجاد امکانات بهداشتی در مناطق محروم و بهبود سیستم ثبت و گزارش‌دهی بیماری‌ها، به‌ویژه آدرس و مشخصات فردی بیماران.

کاهش وسعت منطقه مطالعه و تمرکز بر استان‌ها برای بهبود دقت داده‌ها.

استفاده از مدل شبکه عصبی پرسپترون چندلایه برای مدل‌سازی سرطان‌های دیگر و بررسی عوامل مختلف بیماری. در نظر گرفتن عوامل اجتماعی و شغلی در مدل‌سازی، به‌ویژه برای کشاورزان که در معرض آفتاب زیاد و خطر ابتلا به سرطان پوست هستند.

تهیه نقشه آسیب‌پذیری سرطان پوست برای مناطق مختلف جغرافیایی.

آموزش پیشگیرانه برای اقشار مختلف جامعه، به‌ویژه افرادی که در معرض خطر بالاتری مانند کشاورزان قرار دارند.

- Koch, Stephanie Marie. 2016. *Multilevel Assessment of Skin Cancer Prevention in Arizona*. Thesis, ARIZONA: The University of Arizona.
- Ali-Akbarpour, Mohsen, Abolfazl Mohammadbeigi, Seyed Hamid Reza Tabatabaee, and Gholamreza Hatam. 2012. "Spatial analysis of eco-environmental risk factors of cutaneous leishmaniasis in Southern Iran." *Journal of cutaneous and aesthetic surgery* 30-35.
- Cecconi, Lorenzo, Anna Busolin, Fabio Barbone, Diego Serraino, Alessandra Chiarugi, Annibale Biggeri, and Dolores Catelan. 2016. "Spatial analysis of incidence of cutaneous melanoma in the Friuli Venezia Giulia region in the period 1995-2005." *Geospatial Health*.
- Hagan, Martin T, Howard B Demuth, Mark H Beale, and Orlando De Jesús. 1996. *Neural Network Design (2nd Edition)*. Colorado: Campus Pub. Service.
- Hu, Shasa, Recinda Sherman, Kristopher Arheart, and Robert S. Kirsner. 2014. "Predictors of Neighborhood Risk for Late-Stage Melanoma: Addressing Disparities through Spatial Analysis and Area-Based Measures." *Journal of Investigative Dermatology* 937-945.
- Kotsiantis, S. B. 2007. "Supervised Machine Learning: A Review of Classification Techniques." *Informatica* 249-268.
- Miller, Harvey J., and Jiawei Han. 2009. *Geographic data mining and knowledge discovery*. CRC Press.
- Rajabi, Mohammadreza, Ali Mansourian, and Ahad Bazmani . 2012. "Susceptibility mapping of visceral leishmaniasis based on fuzzy modelling and group decision-making methods." *Geospatial Health* 37-50.
- Rajabi, Mohammadreza, Ali Mansourian, Petter Pilesjö, and Ahad Bazmani. 2014. "Environmental modelling of visceral leishmaniasis by susceptibility-mapping using neural networks: a case study in north-western Iran." *Geospatial Health* 179-191.
- Vega-Corredor, Maria Cecilia, and Jacob Opadeyi. 2014. "Hydrology and public health: linking human leptospirosis and local hydrological dynamics in Trinidad, West Indies." *Earth Perspectives* 1-14.
- Wallingford, Sarah Cristina. 2014. *Investigations into the Epidemiology and Aetiology of Cancers of the Skin*. Thesis(PhD), Manchester: University of Manchester.
- Wheeler, B W, G Kothencz, and A S Pollard. 2013. "Geography of non-melanoma skin cancer and ecological associations with environmental risk factors in England." *British Journal of Cancer* 235-241.
- Zayeri, Farid, Amir Kavousi, and Hadis Najafmehr. 2015. "Spatial analysis of Relative Risks for skin cancer morbidity and mortality in Iran, 2008 – 2010." *Asian Pacific Journal of Cancer Prevention* 5225-5231.
- Alpaydin, Ethem. 2014. *Introduction to machine learning*. Massachusetts: MIT Press.
- Bonita, R. , R. Beaglehole, and T. Kjellström. 2006. *Basic epidemiology 2nd edition*. Switzerland: World Health Organization.

- Brinkhoff, Thomas , Hans-Peter Kriegel, and Bernhard Seeger. 1993. "Efficient processing of spatial joins using R-trees." *ACM* 237-246.
- Brooks, Rodney A. 1991. "Intelligence without representation." *Artificial Intelligence (Artificial Intelligence)* 139-159.
- Buettner, Petra Gertraud , and Robert MacLennan. 2008. "Geographical variation of incidence of cutaneous melanoma in Queensland." *Aust. J. Rural Health* 269–277.
- Burrough, P. A. 1986. *Principles of Geographical Information Systems for Land Resources Assessment*. New York: Oxford University Press.
- Burrough, Peter A. 1986. *Principles of Geographical Information Systems for Land Resources Assessment*. Oxford University Press.
- Hassan, Ali , Hala A Kassem, S. I. El Hosary, and Bahira M El Sawaf. 1999 . "The Use of GIS To Identify Environmental Factors Associated With Cutaneous Leishmaniasis Transmission Risk In Northeast Sinai, Egypt." *Towards Digital Earth*. Beijing: Science Press. 18.
- Hastie, Trevor , Robert Tibshirani, and Jerome Friedman. 2009. *The Elements of Statistical Learning*. New York: Springer .
- Hornik, Kurt. 1989. "Multilayer Feedforward Networks are Universal Approximators." *Neural Networks* 359-366.
- 2008 *How to Write and Illustrate Scientific Papers* Cambridge CAMBRIDGE UNIVERSITY PRESS
- Mohammadinia, Ali , Abbas Alimohammadi, and Rahela Habibi. 2015. "Assessment of environmental factors associated with rural endemics of Leptospirosis in Guilan Province, Iran." *International Conference on Research in Science And Technology*. Kuala Lumpur: researchgate.
- Pezeshki, Z. , M. Tafazzoli-Shadpour, A. Mansourian, B. Eshrati, E. Omid, and I. Nejadqoli. 2012. "Model of cholera dissemination using geographic information systems and fuzzy clustering means: Case study, Chabahar, Iran." *Public Health* 881-887.
- Rätsch, Gunnar. 2004. "A Brief Introduction into Machine Learning."
- Rubens, Philip. 2001. *Science and Technical Writing A Manual of Style*. 2. New York: Routledge.
- Snow, John. 1856. "Cholera and the water supply in the south districts of London in 1854." *Journal of Public Health, and Sanitary Review* 239-57.
- Tatalovich, Zaria , John P. Wilson, Thomas Mack, Ying Yan, and Myles Cockburn. 2006. "The objective assessment of lifetime cumulative ultraviolet exposure for determining melanoma risk." *Journal of Photochemistry and Photobiology B: Biology* 198–204.
- Koch, Stephanie Marie. 2016. *Multilevel Assessment of Skin Cancer Prevention in Arizona*. Thesis, ARIZONA: The University of Arizona.
- Ali-Akbarpour, Mohsen, Abolfazl Mohammadbeigi, Seyed Hamid Reza Tabatabaee, and Gholamreza Hatam. 2012. "Spatial analysis of eco-environmental risk factors of cutaneous leishmaniasis in Southern Iran." *Journal of cutaneous and aesthetic surgery* 30-35.

- Cecconi, Lorenzo, Anna Busolin, Fabio Barbone, Diego Serraino, Alessandra Chiarugi, Annibale Biggeri, and Dolores Catelan. 2016. "Spatial analysis of incidence of cutaneous melanoma in the Friuli Venezia Giulia region in the period 1995-2005." *Geospatial Health*.
- Hagan, Martin T, Howard B Demuth, Mark H Beale, and Orlando De Jesús. 1996. *Neural Network Design (2nd Edition)*. Colorado: Campus Pub. Service.
- Hu, Shasa, Recinda Sherman, Kristopher Arheart, and Robert S. Kirsner. 2014. "Predictors of Neighborhood Risk for Late-Stage Melanoma: Addressing Disparities through Spatial Analysis and Area-Based Measures." *Journal of Investigative Dermatology* 937–945.
- Kotsiantis, S. B. 2007. "Supervised Machine Learning: A Review of Classification Techniques." *Informatica* 249-268.
- Miller, Harvey J., and Jiawei Han. 2009. *Geographic data mining and knowledge discovery*. CRC Press.
- Rajabi, Mohammadreza, Ali Mansourian, and Ahad Bazmani . 2012. "Susceptibility mapping of visceral leishmaniasis based on fuzzy modelling and group decision-making methods." *Geospatial Health* 37-50.
- Rajabi, Mohammadreza, Ali Mansourian, Petter Pilesjö, and Ahad Bazmani. 2014. "Environmental modelling of visceral leishmaniasis by susceptibility-mapping using neural networks: a case study in north-western Iran." *Geospatial Health* 179-191.
- Vega-Corredor, Maria Cecilia, and Jacob Opadeyi. 2014. "Hydrology and public health: linking human leptospirosis and local hydrological dynamics in Trinidad, West Indies." *Earth Perspectives* 1-14.
- Wallingford, Sarah Cristina. 2014. *Investigations into the Epidemiology and Aetiology of Cancers of the Skin*. Thesis(PhD), Manchester: University of Manchester.
- Wheeler, B W, G Kothencz, and A S Pollard. 2013. "Geography of non-melanoma skin cancer and ecological associations with environmental risk factors in England." *British Journal of Cancer* 235–241.
- Zayeri, Farid, Amir Kavousi, and Hadis Najafmehr. 2015. "Spatial analysis of Relative Risks for skin cancer morbidity and mortality in Iran, 2008 – 2010." *Asian Pacific Journal of Cancer Prevention* 5225-5231.
- Alpaydin, Ethem. 2014. *Introduction to machine learning*. Massachusetts: MIT Press.
- Bonita, R. , R. Beaglehole, and T. Kjellström. 2006. *Basic epidemiology 2nd edition*. Switzerland: World Health Organization.
- Brinkhoff, Thomas , Hans-Peter Kriegel, and Bernhard Seeger. 1993. "Efficient processing of spatial joins using R-trees." *ACM* 237-246.
- Brooks, Rodney A. 1991. "Intelligence without representation." *Artificial Intelligence (Artificial Intelligence)* 139-159.
- Buettner, Petra Gertraud , and Robert MacLennan. 2008. "Geographical variation of incidence of cutaneous melanoma in Queensland." *Aust. J. Rural Health* 269–277.
- Burrough, P. A. 1986. *Principles of Geographical Information Systems for Land Resources Assessment*. New York: Oxford University Press.

- Burrough, Peter A. 1986. *Principles of Geographical Information Systems for Land Resources Assessment*. Oxford University Press.
- Hassan, Ali , Hala A Kassem, S. I. El Hosary, and Bahira M El Sawaf. 1999 . "The Use of GIS To Identify Environmental Factors Associated With Cutaneous Leishmaniasis Transmission Risk In Northeast Sinai, Egypt." *Towards Digital Earth*. Beijing: Science Press. 18.
- Hastie, Trevor , Robert Tibshirani, and Jerome Friedman. 2009. *The Elements of Statistical Learning*. New York: Springer .
- Hornik, Kurt. 1989. "Multilayer Feedforward Networks are Universal Approximators." *Neural Networks* 359-366.
- 2008 *How to Write and Illustrate Scientific Papers* Cambridge CAMBRIDGE UNIVERSITY PRESS
- Mohammadinia, Ali , Abbas Alimohammadi, and Rahela Habibi. 2015. "Assessment of environmental factors associated with rural endemics of Leptospirosis in Guilan Province, Iran." *International Conference on Research in Science And Technology*. Kuala Lumpur: researchgate.
- Pezeshki, Z. , M. Tafazzoli-Shadpour, A. Mansourian, B. Eshrati, E. Omid, and I. Nejadqoli. 2012. "Model of cholera dissemination using geographic information systems and fuzzy clustering means: Case study, Chabahar, Iran." *Public Health* 881-887.
- Rätsch, Gunnar. 2004. "A Brief Introduction into Machine Learning."
- Rubens, Philip. 2001. *Science and Technical Writing A Manual of Style*. 2. New York: Routledge.
- Snow, John. 1856. "Cholera and the water supply in the south districts of London in 1854." *Journal of Public Health, and Sanitary Review* 239-57.
- Tatalovich, Zaria , John P. Wilson, Thomas Mack, Ying Yan, and Myles Cockburn. 2006. "The objective assessment of lifetime cumulative ultraviolet exposure for determining melanoma risk." *Journal of Photochemistry and Photobiology B: Biology* 198-204.
- آذرمهر , محمدرضا, محمدسعیدی مسگری , و محمد کریمی . ۱۳۸۹. "مدلسازی مکانی-زمانی بیماری مالاریا با استفاده از Geographical Information System و روش Cellular Automata." فصلنامه بیماری های عفونی و گرمسیری وابسته به انجمن متخصصین بیماری های عفونی و گرمسیری ۶۱-۶۹.
- افضلی , مجید, مسعود میرزایی , حسن سعادت , و سید سعید مظلومی محمودآبادی. ۱۳۹۲. "اپیدمیولوژی سرطان پوست و تغییرات روند زمانی بروز آن در ایران." دوماهنامه علمی-پژوهشی فیض ۵۰۱-۵۱۱.
- العسیری, صفا. ۱۳۹۳. بررسی میزان آگاهی و شناخت بیماران مبتلا به ملانوم و کانسره های غیرملانوم پوست مراجعه کننده به درمانگاه تومور بیمارستان رازی از بیماری خود و عوامل مؤثر بر آن. پایان نامه, تهران: دانشگاه علوم پزشکی و خدمات بهداشتی درمانی تهران.

- آهنگرکانی، مهرداد، مهدی فرنقی، و محمدرضا شیرزادی. ۱۳۹۵. "پیش بینی مکانی-زمانی مناطق پرخطر بیماری لپتوسپیروز با استفاده از روشهای رگرسیون وزندار جغرافیایی و شبکه عصبی پرسپترون چند لایه." *نشریه علمی-پژوهشی علوم و فنون نقشه برداری* ۷۹-۹۸.
- بلیانی، یدالله، و سید یاسر حکیم دو ست. ۱۳۹۳. *اصول و مبانی پردازش دادههای مکانی (فضایی) با استفاده از روش های تحلیل فضایی*. تهران: انتشارات آزادپیما.
- بیاتانی، علی، و علی صادقی. ۱۳۹۱. "تجزیه و تحلیل مکانی عوامل محیطی مؤثر بر لیشمانیوز جلدی در ایران با استفاده از سیستم اطلاعات جغرافیایی." *مجله تحقیقات نظام سلامت حکیم* ۱۵۸-۱۶۵.
- پیروز، الهام. ۱۳۹۰. *بررسی شیوع انواع مختلف ملانوما و شایع ترین ناحیه ی اناتومیکی درگیر در بیماران مراجعه کننده به بیمارستان رازی از فروردین سال ۱۳۸۳ تا فروردین ۱۳۸۸*. پایان نامه، تهران: دانشگاه علوم پزشکی و خدمات بهداشتی درمانی تهران، دانشکده پزشکی.
- ترابی آزاد، مسعود، امیر سیه سرانی، و رحیم افتخاری. ۱۳۹۵. *آموزش جامع تحلیل گر زمین آماری نرم افزار Arc GIS*. تهران: انتشارات سازمان جغرافیایی نیروهای مسلح.
- جعفری، مرتضی، و حسن حبیبی ماچیان. ۱۳۹۳. *راهنمای گام به گام تحلیل داده ها با برنامه spss*. تهران: جامعه شناسان.
- خسروی، یونس، و اسماعیل عباسی. ۱۳۹۵. *تحلیل فضایی داده های محیطی با زمین آمار*. تهران: انتشارات آذر کلک.
- خوان ایرانی، پگاه، محمد کریمی، و علی منصوریان. ۱۳۹۲. "پیشبینی بیماری لیشمانیوز جلدی با استفاده از روش شبکه عصبی مصنوعی مطالعه موردی: استان ایلام." *سی و دومین گردهمایی و نخستین کنگره بینالمللی علوم زمین*. تهران: سازمان زمینشناسی و اکتشافات معدنی کشور.
- رضایی، عاطفه. ۱۳۹۲. *بررسی شیوع ملانوم در نمونه های بیوپسی خال هایی که طی سالهای ۱۳۸۹ تا ۱۳۹۱ به آزماي شگاه آ سیب شنا سی بیمارستان رازی ارجاع شده اند*. پایان نامه، تهران: دانشگاه علوم پزشکی و خدمات بهداشتی درمانی تهران.
- زارع، محمد، پروین شمس زاده، و عباس نجاری. ۱۳۸۵. "فراهم سازی امکان استفاده از سیستم های اطلاعات جغرافیایی در تصمیم گیری های مدیریتی در بخش بهداشت و درمان." *مجله تحقیقات نظام سلامت*.
۱۳۸۵. *سالنامه آماری کشوری*. سالنامه، تهران: مرکز آمار ایران.
- عسگری، علی. ۱۳۹۰. *تحلیل های زمین آماری با ArcGIS*. تهران: انتشارات سازمان فناوری اطلاعات و ارتباطات شهرداری تهران.
- کاتورانی، شلیر، سید علی المدرسی، جلال کرمی، و رضا ریسی. ۱۳۹۱. "تحلیل فضایی بیماری سرطان پو ست در استان کردستان در محیط GIS." *دومین سمینار تخصصی کاربردهای علوم و فناوری های ژئوانفورماتیک در نظام سلامت (ژئوانفورماتیک در خدمت محیطی ست سالم)*. تهران: موسسه ملی تحقیقات سلامت جمهوری اسلامی ایران و پژوهشکده ژئوانفورماتیک.

- کاظمی, آفاق, امید اسکندری, و محمد کریمی. ۱۳۹۴. "بروز و توزیع جغرافیایی سرطان پوست در استان کردستان ایران." پوست و زیبایی ۳۸-۴۵.
- کیا, مصطفی. ۱۳۸۷. شبکه های عصبی در *MATLAB*. تهران: کیان رایانه سبز.
۱۳۸۹. گزارش کشوری ثبت موارد سرطان سال ۱۳۸۲-۱۳۸۹. گزارش, تهران: وزارت بهداشت, معاونت بهداشت و درمان, مرکز مدیریت بیماریهای غیر واگیر, اداره سرطان.
۱۳۸۹. گزارش کشوری ثبت موارد سرطان سال ۱۳۸۷. گزارش, تهران: وزارت بهداشت, معاونت بهداشت و درمان, مرکز مدیریت بیماریهای غیر واگیر, اداره سرطان.
- مبارک آبادی, عادل, راضیه رجبی, یوسف خانی, و امیر الماسی حشینی. ۱۳۹۲. "همه گیرشناسی سرطان پوست در استان مرکزی, ایران." پوست و زیبایی ۱۲۰-۱۲۶.
- محمدی نیا, علی. ۱۳۹۴. مقایسه رگرسیون وزن دار جغرافیایی, شبکه عصبی مصنوعی و الگوریتم ژنتیک در مدل سازی مکانی تب شالیزار. پایان نامه, تهران: دانشگاه خواجه نصیرالدین طوسی.
- معصومی, زهزه. ۱۳۸۴. به کارگیری سیستم اطلاعات جغرافیایی در پایش و مدیریت بیماری ها. پایان نامه, تهران: دانشگاه صنعتی خواجه نصیرالدین طوسی, ۷۵۳-۷۶۳.
- ملالو, ابوالفضل. ۱۳۹۲. تحلیل و مدلسازی مکانی- زمانی بیماری سالک. پایان نامه, تهران: دانشگاه خواجه نصیرالدین طوسی.
- منهاج, محمدباقر. ۱۳۹۵. مبانی شبکه های عصبی: هوش محاسباتی. تهران: دانشگاه صنعتی امیرکبیر.
۱۳۹۶. مؤسسه تحقیقات, درمان و آموزش سرطان ایران. ۲۰۴. دستیابی در ۲۰۴, ۱۳۹۶. <http://ncii.ir>
- نوحه گر, احمد, و محمد اکبری نسب. ۱۳۹۲. تجزیه و تحلیل فضایی *Arc GIS 9.X* هرمزگان: دانشگاه هرمزگان.
- نورعلی نژاد, مایسا. ۱۳۹۳. بررسی مشخصات بیماران مبتلا به ملانوم مراجعه کننده به بیمارستان رازی تهران از فروردین ۱۳۸۳ تا فروردین ۱۳۹۳. پایان نامه, تهران: دانشگاه علوم پزشکی و خدمات بهداشتی درمانی تهران.