

ارائه مدلی برای پیش‌بینی بیماری لیشمانیوز جلدی (سالک) با استفاده از سامانه اطلاعات مکانی و الگوریتم شبکه عصبی

مهرداد آهنگرکانی^۱

مهدی فرنقی^۲

تاریخ دریافت مقاله: ۹۶/۰۶/۰۹

تاریخ پذیرش مقاله: ۹۶/۱۰/۰۳

چکیده

بیماری سالک، از بیماری‌های انگلی می‌باشد که در شمار بیماری‌های مشترک بین انسان و حیوان قرار می‌گیرد. این بیماری از شایع‌ترین فرم بیماری لیشمانیوز است که توسط گونه‌های مختلف انگل لیشمانیا ایجاد شده و با نیش زدن گونه‌های مختلف پشه خاکی‌های ماده عامل فلپوتومینه به انسان، شخص را دچار ابتلا به این بیماری می‌کند. استان گلستان همواره یکی از کانون‌های اصلی بروز بیماری سالک در ایران بوده است و به دلیل دارا بودن شرایط محیطی و آب و هوایی مساعد، سالانه تعدادی از موارد ابتلا به این بیماری در این استان گزارش می‌گردد. هدف اساسی این تحقیق تحلیل سالانه توزیع مکانی-زمانی بیماری سالک، بررسی تأثیر عوامل محیطی و آب و هوایی با بروز بیماری و در نهایت ارائه مدلی جهت تهیه نقشه پیش‌بینی و آسیب‌پذیری بیماری طی دوره آماری ۱۳۹۲ تا ۱۳۹۴ در سطح دهستان‌های استان گلستان می‌باشد. به منظور بررسی ارتباط میان بروز بیماری سالک با متغیرهای محیطی و آب و هوایی و همچنین بررسی وجود خودهمبستگی مکانی میان موارد بروز بیماری، تحلیل‌های آماری و مکان-آماری به کار گرفته شده‌اند. جهت مدل‌سازی بیماری، الگوریتم شبکه عصبی پرسپترون چندلایه مورد استفاده قرار گرفت. به منظور ارزیابی دقت مدل بدست آمده، معیارهایی همچون $RMSE$ ، $MAPE$ و AUC استفاده گردیدند و همچنین جهت تعیین مؤثرترین متغیرها در مدل‌سازی بیماری، آنالیز حساسیت اجرا شده است. معیارهای ارزیابی گویای این حقیقت بودند که مدل به دست آمده قدرت تشخیص قابل قبولی در پیش‌بینی بروز بیماری در سطح دهستان‌های استان گلستان دارد ($RMSE_{1392} = 0.019$, $RMSE_{1393} = 0.013$, $RMSE_{1394} = 0.017$, $MAPE_{1392} = 1.43$, $MAPE_{1393} = 1.34$, $MAPE_{1394} = 1.40$, $AUC_{1392} = 0.846$, $AUC_{1393} = 0.873$, $AUC_{1394} = 0.859$). همچنین آنالیز حساسیت نشان داد که متغیرهای پوشش گیاهی و متوسط رطوبت هوا مهم‌ترین عوامل در تهیه نقشه پیش‌بینی و آسیب‌پذیری توزیع مکانی بیماری سالک در استان گلستان می‌باشند.

واژه‌های کلیدی: بیماری سالک، سامانه اطلاعات مکانی، شبکه عصبی پرسپترون چندلایه، آنالیز حساسیت، گلستان

۱- دانشجوی دکتری سامانه اطلاعات مکانی، دانشگاه صنعتی خواجه نصیرالدین طوسی (نویسنده مسئول) mahangar@mail.kntu.ac.ir

۲- استادیار گروه سامانه اطلاعات مکانی دانشکده نقشه‌برداری، دانشگاه صنعتی خواجه‌نصیرالدین طوسی farnaghi@kntu.ac.ir

۱- مقدمه

دیده می‌شود. کشور ایران با بالاترین میزان بروز سالک روستایی در شمار کشورهای است که عفونت با لیشمانیا ماژور^۵ در اغلب مناطق آن به شکل آندمیک مشاهده می‌شود (Rassi, Javadian, Jalali, Motazedian, & Vatndoost, 2004). موارد زیادی از بیماری سالک از استان گلستان گزارش می‌شود و این استان از کانون‌های مهم بیماری لیشمانیوز جلدی روستایی ایران محسوب می‌شود. در اپیدمیولوژی بیماری لیشمانیوز مشخص شده است که شیوع و انتشار این بیماری تحت تأثیر مسائل اقتصادی، اجتماعی، فرهنگی و به ویژه شرایط محیطی و آب و هوایی همچون درجه حرارت، رطوبت نسبی، بارندگی، پوشش گیاهی است که از مهمترین عوامل محیطی و آب و هوایی در بروز بیماری سالک محسوب می‌شوند (Organization, 2009).

انتشار بیماری یک پدیده دینامیکی است که در مکان و با گذشت زمان تغییر می‌کند. مسائل بهداشت عمومی و انتشار بیماری‌ها با جغرافیای منطقه رابطه مستقیم دارند و همواره دارای بُعد مکانی هستند و توزیع و تمرکز آنها در مکان‌های مختلف، متفاوت است (AhangarCani, Farnaghi, Shirzadi, 2016; Massoomy & Mesgari, 2006). تمرکز یک بیماری در مناطقی خاص، بروز غیرمعمول و وجود یک منبع ایجادکننده بیماری را نشان می‌دهد. از این رو مطالعه جغرافیای منطقه در بررسی بیماری‌ها از اهمیت خاصی برخوردار است (AhangarCani et al., 2016; Hassan, Kenawy, Kamal, Abdel Sattar, & Sowilem, 2003). دسترسی به اطلاعات مکانی مفید و اتخاذ تصمیمات مناسب بر مبنای تحلیل چنین اطلاعاتی موجبات ارتقای سطح سلامت و ایجاد جامعه‌ای بهره‌مند از سلامت جسمی و روحی در زمان کوتاه‌تر و با هزینه کمتر را فراهم خواهد نمود. غالب سازمان‌های رفاهی به ابزارهایی نیاز دارند که به پیش، کنترل و مدیریت بیماری‌ها کمک کنند (McLafferty, 2003). در سال‌های اخیر سامانه‌های اطلاعات مکانی^۶ تحولات

مهمترین اهداف هر سازمان بهداشتی، سالم‌سازی محیط زیست، کنترل بیماری‌ها، آموزش بهداشت، پیشگیری از بروز بیماری، اقدامات پزشکی و پرستاری جهت تشخیص زودرس، کنترل و مدیریت بیماری‌ها می‌باشد (Bayatani & Sadeghi, 2012). آشنایی با بیماری‌ها و خطراتی که سلامت جامعه را تهدید می‌کند بدون آشنایی و استفاده از دانش و روش‌های همه گیر شناسی جهت تعیین عوامل خطر و چگونگی برخورد با آنها کمتر عملی است (Gordis, 2000). لیشمانیوز^۱ نام گروهی از بیماری‌های انگلی می‌باشد که به اشکال گوناگونی از جمله احشایی^۲، جلدی^۳ (سالک) و مخاطی^۴ ظاهر شده و در شمار بیماری‌های مشترک بین انسان و حیوان قرار می‌گیرد (A.MOLLALO, 2014; Mollalo, Alimohammadi, Shirzadi, & Malek, 2015). بیماری سالک شایع‌ترین فرم بیماری لیشمانیوز می‌باشد که توسط گونه‌های مختلف انگل لیشمانیا ایجاد شده و با نیش زدن گونه‌های مختلف پشه خاکی‌های ماده عامل فلبوتومینه به انسان، شخص را دچار ابتلا به این بیماری می‌کند (A.MOLLALO, 2014). با وجود پیشرفت‌های روز افزون بشر در زمینه کنترل بیماری‌ها، هنوز هم لیشمانیوزها یکی از معضلات و مشکلات مهم بهداشتی- اقتصادی بسیاری از کشورهای جهان از جمله ایران می‌باشند (A.MOLLALO, 2014; Assimina, Charilaos, & Fotoula, 2008). بسیار دور در ایران وجود داشته و امروزه کشور ایران یکی از کانون‌های مهم بیماری سالک در جهان محسوب می‌شود. بیماری سالک به دو نوع شهری و روستایی تقسیم می‌شود. در ایران نوع شهری بیماری سالک در تهران، مشهد، شیراز، کرمان، زاهدان، یزد و بم به صورت بومی وجود دارد و نوع روستایی آن در قسمت‌های شمال و شمال شرق گلستان، اصفهان، خوزستان، گنبد کاووس، اسفراین، بوشهر و فارس

1- Leishmaniasis

2- Viscera Leishmaniasis

3- Cutaneous Leishmaniasis

4- Mucocutaneous Leishmaniasis

5- Leishmania Major

6- Geospatial Information System (GIS)

مکانی، زمانی و مکانی- زمانی بیماری لیشمانیوز احشایی را در مناطق مرکزی برزیل تحلیل نمودند و به این نتیجه رسیدند که مسیر گسترش و انتشار بیماری در جهت شرق به غرب و همزمان با مداخلات انسانی از قبیل ساخت بزرگراه، راه آهن و لوله گاز بوده است (Antonialli, Torres, Paranhos Filho, & Tolezano, 2007). همکاران در سال ۲۰۰۶ میلادی برای بررسی بیماری تب شالیزار در ایالت فرا تایلند در فاصله زمانی ۱۹۹۸ تا ۲۰۰۴ میلادی از ترکیب سامانه اطلاعات مکانی و سنجش از دور بهره بردند. برای استخراج متغیرهای شاخص‌های پوشش گیاهی NDVI و SAVI و کاربری اراضی از تصاویر ماهواره‌ای ASTER استفاده نمودند. نتایج به دست آمده گویای این حقیقت بوده است که بیشترین نرخ بروز بیماری مربوط به مناطق پست با پوشش گیاهی فراوان و مناطق روستایی می‌باشد (Herbreteau et al., 2006).

اما در خصوص بیماری سالک، تاکنون در سرتاسر جهان مطالعات مختلفی انجام شده است که در این خصوص می‌توان اشاره نمود که در بررسی انجام شده در برزیل در سال ۲۰۰۹ میلادی در مورد بیماری سالک با استفاده از سنجش از دور به منظور بررسی انواع پوشش گیاهی و مکان‌های با وقوع بالای بیماری به این نتیجه رسیدند که بیشترین موارد بیماری در سواحل رودخانه‌ها و دریاچه‌های با پوشش گیاهی درختی اتفاق افتاده‌اند (Soccol et al., 2009). همچنین مورون و همکاران در سال ۲۰۱۱ میلادی، با بررسی خصوصیات اپیدمیولوژی و جغرافیایی سالک بیان نمودند که ویژگی‌های محیطی و فردی همچون جنسیت مذکر، فعالیت در فضای آزاد، سطح تحصیلات پایین و زندگی در مناطق با ارتفاع زیاد، بیشترین تأثیر را در خطر ابتلاء به بیماری سالک دارند (Morrone et al., 2011).

در پژوهش چاوز و پاسکول ارتباط معنی‌داری بین شرایط آب و هوایی و بروز بیماری لیشمانیوز جلدی مشاهده گردید. همچنین این محققین عنوان نمودند که بسته به زمان پیشگویی می‌توان به میزان ۷۷-۷۲ درصد

و مدیریت داده‌های مکانی ایجاد کرده است. این علم و فناوری با داشتن خصوصیتی مانند قابلیت اخذ و تبادل داده‌ها از منابع مختلف، سازماندهی اطلاعات، دریافت و نمایش به هنگام اطلاعات، پردازش و تلفیق داده‌ها، افق‌های جدیدی را بر روی محققان گشوده است (Zare, Shamszadeh, & Najjari, 2006). بنابراین سامانه‌های اطلاعات مکانی می‌توانند به عنوان یک علم و تکنولوژی راهگشا، وارد مسائل بهداشت عمومی شوند. در نهایت می‌توان اشاره نمود که با توجه به ارتباط مستقیم انتشار بیماری و مسائل بهداشت عمومی با جغرافیای منطقه، سامانه‌های اطلاعات مکانی با توجه به دارا بودن قابلیت‌های فراوان در این زمینه، نقش بسزایی در مدیریت و برنامه‌ریزی مسائل بهداشت عمومی و مطالعه انتشار بیماری‌ها ایفا می‌نمایند. (AhangarCani et al., 2016; Mesgari & Masoomi, 2008).

مطالعات فراوانی به منظور بررسی بیماری‌های مختلف در سراسر جهان صورت گرفته‌اند و روش‌های بسیاری با استفاده از علوم پزشکی و مهندسی در این زمینه به کار گرفته شده‌اند که بیشتر در مورد بیماری‌های واگیر بوده‌اند. امروزه سامانه اطلاعات مکانی به عنوان بخشی از فناوری اطلاعات، توسط سازمان‌های مدیریت بهداشت و درمان، در سطوح مختلف جهانی، ملی، منطقه‌ای و محلی به طور گسترده مورد استفاده قرار گرفته است (Zare et al., 2006). در سطح بین‌المللی مطالعات گسترده‌ای با کاربرد تکنیک‌های سامانه اطلاعات مکانی، سنجش از دور و تحلیل‌های مکان-آماری متعدد، در مورد بیماری‌های واگیر مشترک میان انسان و حیوان گزارش شده است. مطالعه انجام شده در برزیل با استفاده از سامانه اطلاعات مکانی و سنجش از دور در مورد بیماری لیشمانیوز احشایی نشان داده است که پوشش گیاهی یکی از مهمترین عوامل محیطی مؤثر بر شیوع این بیماری واگیر بوده است. بدین صورت که بیشتر ناقلین در مطالعه فوق در مناطق دارای پوشش گیاهی کم مشاهده شده‌اند (Bavia, Carneiro, da Costa Gurgel, Filho, & Barbosa, 2005). همچنین آنتونیالی و همکاران در سال ۲۰۰۷ میلادی، توزیع

نشان داد که این بیماری عمدتاً در بخش‌های مرکزی ایران با آب و هوای خشک و بیابانی رخ داده است. همچنین در این دوره زمانی، خوشه‌های بیماری غالباً در استان‌های یزد، خوزستان و کهگیلویه و بویر احمد واقع شده‌اند (Holakouie-Naieni, Mostafavi, Bolorani, Mohebbali, & Pakzad, 2017). در سال‌های اخیر مطالعات زیادی در سرتاسر جهان از روش شبکه‌های عصبی مصنوعی برای مسائل طبقه‌بندی و پیش‌بینی استفاده نموده‌اند. این مطالعات نشان دادند که روش شبکه‌های عصبی مصنوعی به دلیل توانایی بالا در حل مسائل غیر خطی و پیچیده، می‌تواند روش‌های مناسبی به منظور شناسایی، طبقه‌بندی و پیش‌بینی توزیع پدیده‌های دنیای واقعی از جمله بیماری محسوب شوند (Rajabi, Mansourian, 2014; Pilesjö, & Bazmani, 2014; Yilmaz, 2009). از میان مطالعات انجام شده در زمینه بیماری با استفاده از روش شبکه‌های عصبی مصنوعی می‌توان به مطالعات زیر اشاره نمود:

رئوفی و همکاران در سال ۲۰۱۱ میلادی، روش شبکه عصبی مصنوعی را در شناسایی بیماری هیپاتیت مورد بررسی قرار دادند (Raoufy et al., 2011). همچنین رجبی و همکاران در سال ۲۰۱۴ میلادی، از مدل شبکه عصبی RBFLN جهت تهیه نقشه آسیب‌پذیری بیماری لیشمانیوز احشایی در ایران استفاده نمودند و به این نتیجه دست یافتند که حاشیه رودخانه، مناطق کم ارتفاع و با پوشش گیاهی کم و پراکنده و روستاهای عشایرنشین بدون مراکز درمانی و بهداشتی واقع در منطقه مورد مطالعه، مکان‌های مناسبی برای رشد و تکثیر مخازن و ناقلین بیماری محسوب می‌شوند (Rajabi et al., 2014).

تمامی مطالعات انجام شده در رابطه با بیماری سالک، به بررسی ارتباط میان وقوع این بیماری با عوامل مؤثرش تمرکز کرده‌اند. از این‌رو، با توجه به بررسی‌های انجام شده و طبق بهترین اطلاعات، در هیچ کدام از مطالعات پیشین، توزیع مکانی- زمانی بیماری سالک بررسی و ارزیابی نشده و همچنین راهکاری جهت تهیه نقشه پیش‌بینی و آسیب‌پذیری توزیع مکانی بیماری سالک در استان بومی گلستان به منظور فراهم نمودن اطلاعات لازم جهت

روند گسترش بیماری را به صورت صحیح پیش‌بینی نمود (Chaves & Pascual, 2006). رجبی و همکاران در سال ۲۰۱۶ میلادی در اصفهان، مدلی عامل مبنای به منظور شبیه‌سازی پویای گسترش بیماری سالک براساس یک سیستم اتوماتای جغرافیایی^۱ توسعه دادند. نتایج این تحقیق نشان دادند که مناطق بیابانی منشأ اصلی بیماری سالک می‌باشند و جمعیت روستایی در مواجهه بیشتری با ناقلین و مخازن بیماری قرار دارند (Rajabi, Pilesjö, Shirzadi, Fadaei, & Mansourian, 2016). مظفری و همکاران در سال ۱۳۹۰ به بررسی رابطه میان پوشش گیاهی و میزان شیوع بیماری سالک با استفاده از سامانه اطلاعات مکانی در سطح دشت یزد- اردکان پرداختند و به این نتیجه رسیدند که عامل پوشش گیاهی به طور مستقیم و غیر مستقیم نشانگر پارامترهای مؤثر بر بیماری سالک است و بیشترین میزان بیماری در سطح منطقه مورد مطالعه در مناطقی با پائین‌ترین میزان پوشش گیاهی متمرکز شده است (Mozaffari, BAKHSHIZADEH, & GHEIBI, 2012). همچنین رمضان خانی در سال ۲۰۱۷ میلادی به بررسی ارتباط میان بروز بیماری سالک با عوامل محیطی همچون دما، بارش، رطوبت، سرعت باد، پوشش گیاهی، ارتفاع و تراکم جمعیت با استفاده از روش‌های رگرسیون وزن‌دار جغرافیایی (GWR)^۲ و رگرسیون کمترین مربعات معمولی (OLS)^۳ در استان اصفهان پرداختند. در این تحقیق محققان به این نتیجه رسیدند که روش GWR بهتر توانسته است ارتباط بین عوامل با بروز بیماری سالک را در مقایسه با روش OLS مدل‌سازی کند. همچنین نتایج این تحقیق نشان داده است که همبستگی مثبت میان دما و تراکم جمعیت با بروز بیماری وجود دارد (Ramezankhani, Hosseini, Sajjadi, Khoshabi, & Ramezankhani, 2017). نایینی و همکاران در سال ۲۰۱۷، به بررسی توزیع مکانی بیماری سالک طی دوره زمانی ۲۰۱۳-۱۹۸۳ در ایران پرداختند. نتایج این تحقیق

1- Agent-Based Model (ABM)

2- Geographic Automata System (GAS)

3- Geographically Weighted Regression (GWR)

4- Ordinary Least Square (OLS)

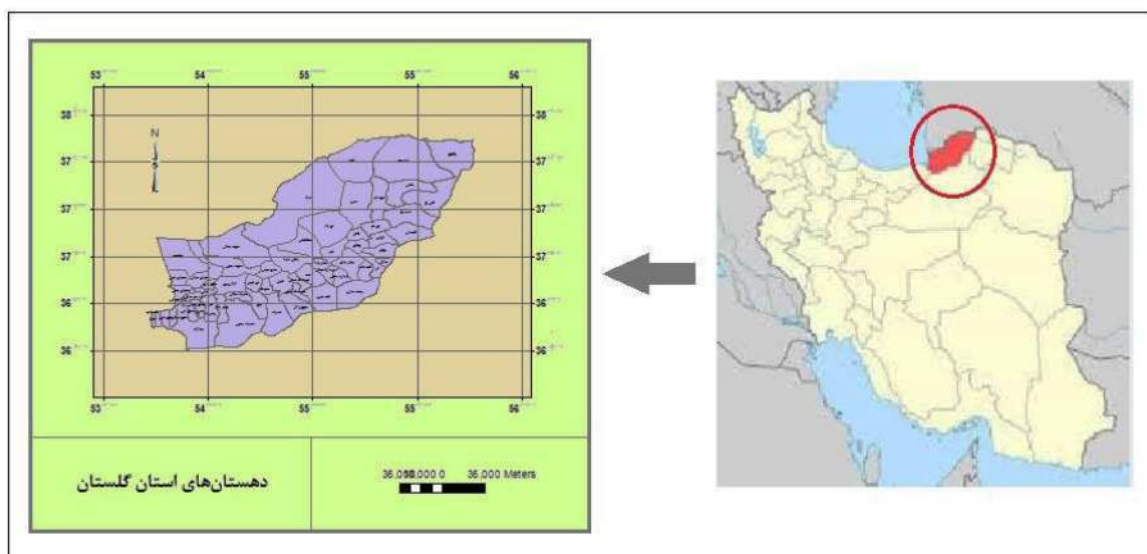
فصلنامه علمی - پژوهشی اطلاعات جغرافیایی (ص ۱۱)
ارائه مدلی برای پیش‌بینی بیماری لیشمانیوز جلدی (سالک) ... / ۱۱

و ارزیابی آنها و در نهایت آنالیز حساسیت ارائه و بحث شده‌اند. در فصل پایانی، به جمع‌بندی و نتیجه‌گیری مطالب ارائه شده در فصول قبلی و پیشنهادات و راهکارهایی برای مطالعات آینده پرداخته شده است.

۲- معرفی محدوده مورد مطالعه و داده‌های مورد استفاده

با توجه به بومی بودن بیماری سالک در استان گلستان و شیوع بالای این بیماری در این استان به دلیل دارا بودن شرایط محیطی و آب و هوایی مناسب جهت رشد و تکثیر مخازن و ناقلین بیماری، در این تحقیق، استان گلستان به عنوان منطقه مورد مطالعه انتخاب شده است. استان گلستان از نظر تقسیمات سیاسی متشکل از ۶۰ دهستان و مربوط به زون ۴۰ شمالی سیستم تصویر UTM با طول جغرافیایی ۵۳ درجه و ۵۷ دقیقه تا ۵۶ درجه و ۲۲ دقیقه شرقی و عرض جغرافیایی ۳۶ درجه و ۳۰ دقیقه تا ۳۸ درجه و ۸ دقیقه شمالی واقع در شمال شرقی ایران می‌باشد (نگاره ۱). طبق آمار ارائه شده از سرشماری مرکز بهداشت استان گلستان در سال ۱۳۹۰، این استان دارای ۱۷۶۷۳۴۵ نفر جمعیت بوده است که ۹۰۲۶۸۸ نفر آن مربوط به جمعیت شهری و

برنامه‌ریزی و سیاست‌گذاری با هدف ارتقای سطح سلامت و بهداشت عمومی ارائه نشده است. بنابراین، اهداف اصلی این تحقیق، تحلیل سالانه توزیع مکانی-زمانی بیماری سالک، بررسی تأثیر عوامل محیطی و آب و هوایی با بروز بیماری و در نهایت ارائه مدلی جهت تهیه نقشه پیش‌بینی و آسیب‌پذیری بیماری سالک طی دوره آماری ۱۳۹۲ تا ۱۳۹۴ در سطح دهستان‌های استان گلستان می‌باشد. بر این اساس مقاله حاضر در ۵ فصل به قرار زیر تدوین شده است. در فصل اول مقدمه‌ای در رابطه با بیماری‌ها به خصوص بیماری‌های واگیر و مشترک میان انسان و حیوان، نقش سامانه اطلاعات مکانی در مدل‌سازی بیماری‌ها و مطالعات صورت گرفته در سراسر جهان در رابطه با بیماری‌های واگیر به طور خاص بیماری سالک ارائه شده است. در فصل دوم منطقه مورد مطالعه و داده‌های مورد استفاده در این تحقیق تشریح شده‌اند. روش‌های آماری و تحلیلی به همراه روش مورد استفاده جهت مدل‌سازی بیماری سالک در قالب مبانی نظری تحقیق در فصل بعدی توضیح داده شده‌اند. در فصل چهارم، نحوه تهیه نقشه پیش‌بینی و آسیب‌پذیری بیماری سالک و همچنین نتایج حاصل از روش‌های آماری و تحلیلی به همراه نقشه‌های پیش‌بینی



نگاره ۱: دهستان‌های مربوط به استان گلستان به عنوان محدوده مورد مطالعه

دیگر مرتبط است اما چیزهایی که بهم نزدیکترند بیشترین تأثیر را نسبت به چیزهای دورتر بر همدیگر دارند. (Tobler, 1970)، به نظر می‌رسد که خود همبستگی مکانی میان موارد بروز بیماری سالک ناشی از موقعیت مطلق آنها و همچنین موقعیت هر یک از آنها نسبت به موارد دیگر باشد. بنابراین روش‌های آماری معمول نمی‌توانند در تحلیل چنین داده‌هایی مورد استفاده قرار گیرند. در این تحقیق به منظور بررسی وجود خودهمبستگی مکانی میان موارد بروز بیماری سالک و تعیین الگوی توزیع مکانی بیماری در منطقه مورد مطالعه از شاخص عمومی Global Moran, s I به دلیل کارایی بالا در داده‌های چندضلعی (AhangarCani et al., 2016; Ruiz, Tedesco, McTighe, Austin, & Kitron, 2004) استفاده شده است.

۸۶۴۶۵۷ نفر آن در روستاها ساکن بوده‌اند. در این تحقیق، داده‌های مربوط به وقوع بیماری سالک، داده‌های جمعیتی و سرشماری و داده‌های سنجش از دوری و آب و هوایی مطابق با جدول (۱) جمع‌آوری شده‌اند. بعد از جمع‌آوری داده‌ها، لایه‌های اطلاعاتی هر یک از آنها به منظور استفاده در تحلیل‌های مورد نیاز آتی، به فرمت شیپ فایل در نرم افزار Arc GIS 10.2.2 تولید شدند.

۳- مبانی نظری

در این قسمت از تحقیق، مفاهیم پایه تحلیل‌های آماری و مکانی و همچنین روش مورد استفاده جهت مدل‌سازی بیماری سالک تشریح می‌شوند.

۳-۱-۱- شاخص عمومی Global Moran, s I

با استفاده از شاخص عمومی Global Moran, s I می‌توان درجه پراکنده بودن یا متمرکز بودن داده‌های مکانی را در فضا اندازه‌گیری نمود. مقادیر شاخص موران بین -۱ تا +۱ متغیر است. به طور کلی، مقادیر مثبت، بیانگر خودهمبستگی مثبت و مقادیر منفی بیانگر خودهمبستگی منفی می‌باشند.

۳-۱- تحلیل‌های مکان-آمار

تحلیل‌های مکان-آمار، مجموعه‌ای از روش‌های آماری هستند که در آنها مکان نقش اصلی را در بررسی داده‌های مرتبط با پدیده‌های مکانی و مکانی-زمانی دارا می‌باشد (Ribeiro Jr & Diggle, 2001; Ripley, 1981). مطابق با قانون اول توبلر در جغرافیا که بیان می‌کند هر چیزی به هر چیز

جدول ۱: داده‌های جمع‌آوری شده جهت مدل‌سازی بیماری سالک

نقشه مبنا	منبع	داده
$= \frac{\text{نرخ بروز استاندارد بیماری سالک} \times \text{تعداد افراد مبتلا در یک جمعیت در یک زمان مشخص}}{\text{تعداد کل جمعیت در همان زمان}}$	مرکز مدیریت بیماری‌های واگیر استان گلستان	بیماری سالک
	مرکز آمار ایران	جمعیتی و سرشماری
شاخص پوشش گیاهی نرمال‌شده ^۱	تصویر ماهواره‌های MODIS با ابعاد پیکسل ۲۵۰ متر	سنجش از دور
متوسط رطوبت هوا متوسط دمای هوا مجموع بارش سالیانه تعداد روزهای بارانی تعداد روزهای یخبندان حداکثر سرعت باد میزان تبخیر	سازمان هواشناسی ایران	هواشناسی

1-Normalized Difference Vegetation Index(NDVI)

فصلنامه علمی - پژوهشی اطلاعات جغرافیایی (مهر)
 ارائه مدلی برای پیش‌بینی بیماری لیشمانیوز جلدی (سالک) /... / ۱۳

که در آن x_i و x_j به ترتیب تعداد موارد بیماری در i امین و j امین واحد مکانی (دهستان) می‌باشند. \bar{x} ، بیانگر متوسط تعداد موارد بیماری در هر دهستان می‌باشد. N تعداد دهستان‌ها در استان گلستان را نشان می‌دهد و W_{ij} نشان دهنده وزن مکانی میان موارد i و j می‌باشد.

۲-۳- تحلیل‌های آماری

تحلیل‌های آماری مجموعه‌ای از روش‌های آماری می‌باشند که به بررسی ارتباط میان متغیرهای مستقل و متغیر وابسته می‌پردازند. در این تحقیق، به منظور بررسی ارتباط میان متغیرهای محیطی و آب و هوایی مؤثر بر بروز بیماری سالک از تحلیل‌های آماری Pearson و Spearman استفاده شده است. در ادامه هر یک از این روش‌ها تشریح می‌شوند.

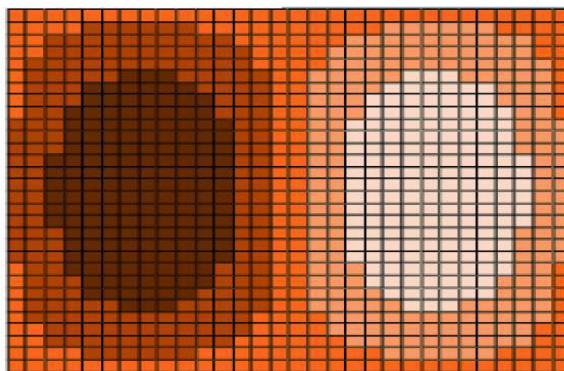
۱-۲-۳- تحلیل همبستگی Pearson

ضریب همبستگی Pearson رابطه مستقیم یا معکوس و همچنین شدت رابطه بین دو یا چند متغیر کمی و پیوسته را نشان می‌دهد. مقدار ضریب همبستگی (r) همیشه بین -1 و 1 می‌باشد. اگر مقدار ضریب برابر -1 باشد در آن صورت همبستگی منفی و کامل می‌باشد. اگر مقدار ضریب برابر 1 باشد در آن صورت همبستگی مثبت و کامل می‌باشد و اگر چنانچه مقدار ضریب برابر 0 باشد نشانه عدم وجود همبستگی میان متغیرهاست. به منظور محاسبه ضریب همبستگی nPearso از رابطه (۲) استفاده می‌شود.

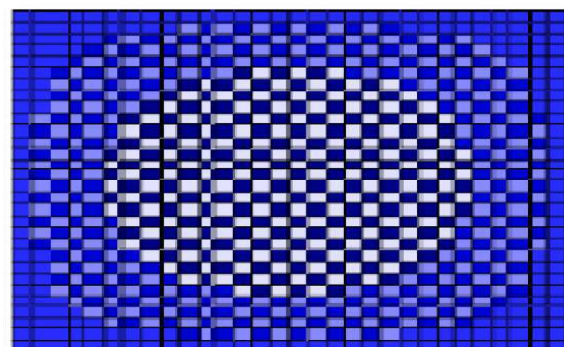
$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2)$$

در رابطه (۲)، r مقدار همبستگی بین متغیرهای x و y ، n : تعداد آزمودنی‌ها (دهستان‌ها)، x_i و y_i : به ترتیب بیانگر مقادیر متغیرهای x و y در i امین دهستان، \bar{x} و \bar{y} به ترتیب متوسط مقادیر متغیرهای x و y می‌باشند. در این

چنانچه مقدار شاخص موران برابر صفر باشد نشان‌دهنده عدم وجود خودهمبستگی مکانی می‌باشد. مطابق با نگاره‌های (۲) و (۳)، اگر مقدار شاخص به عدد $+1$ نزدیک باشد، داده‌ها دارای الگوی خوشه‌ای و اگر مقدار شاخص به عدد -1 نزدیک باشد، داده‌ها از هم گسسته و پراکنده هستند (Moran, 1950).



نگاره ۲: خودهمبستگی مکانی مثبت (توزیع خوشه‌ای)
 (Mitchell, 2005)



نگاره ۳: خودهمبستگی مکانی منفی (توزیع پراکنده)
 (Mitchell, 2005)

مقدار شاخص عمومی Moran's I را می‌توان با استفاده از رابطه (۱) محاسبه نمود (Mitchell, 2005):

$$I = \frac{N}{\sum_{i=1}^N \sum_{j=1}^N W_{ij}} * \frac{\sum_{i=1}^N \sum_{j=1}^N W_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2} \quad (1)$$

مشخص عمل می‌کنند. از مزیت شبکه عصبی نسبت به روش‌های دیگر همچون رگرسیون، این است که در یک شبکه عصبی، بر اساس روابط ذاتی میان داده‌ها، مدلی خطی و یا غیر خطی بین متغیرهای مستقل و وابسته برقرار می‌شود و در این روش به منظور پیش‌بینی و مدل‌سازی یک الگو، نیازی به یک مدل اولیه در برقراری ارتباط میان داده‌های ورودی و خروجی نمی‌باشد. (AhangarCani et al., 2016; Mohammady & Delavar, 2014; Skapura, 1996) روش‌های شبکه عصبی که قادرند برای حل مسائل مختلف مورد استفاده قرار گیرند، پرسپترون‌های چندلایه هستند که توانایی حل مسائل غیر خطی و پیچیده را دارند (Hornik, Stinchcombe, & White, 1989)

در یک شبکه عصبی پرسپترون چندلایه، هر نرون در هر لایه به تمامی نرون‌های لایه قبل متصل می‌باشد. از اینرو به چنین شبکه‌هایی شبکه‌های کاملاً مرتبط نیز می‌گویند. شبکه عصبی پرسپترون چندلایه از یک لایه ورودی، یک یا چند لایه میانی و یک لایه خروجی تشکیل شده است. به منظور آموزش شبکه‌های عصبی پرسپترون چندلایه با هدف تنظیم پارامترهای شبکه و نزدیکی پاسخ شبکه به پاسخ واقعی، معمولاً از قانون یادگیری پس انتشار خطا^۳ استفاده می‌شود. برای کسب اطلاعات بیشتر در رابطه قانون یادگیری پس انتشار خطا، به Williams and Hinton (۱۹۸۶) رجوع شود.

۴- اجرا، نتایج و بحث

در این قسمت، علاوه بر ارائه روند کلی تحقیق (نگاره ۴)، نحوه مدل‌سازی بیماری سالک به منظور تهیه نقشه‌های پیش‌بینی و آسیب‌پذیری بیماری به همراه نتایج حاصل از تحلیل‌های مکان-آمار، نقشه‌های پیش‌بینی و ارزیابی آنها با استفاده از معیارهای ارزیابی و نتایج آنالیز حساسیت ارائه و بحث شده‌اند.

تحقیق تحلیل همبستگی Pearson به منظور بررسی ارتباط میان متغیرهای محیطی و آب و هوایی مؤثر و نوع و شدت ارتباط هر یک از این متغیرها بر بروز بیماری سالک در سطح استان گلستان استفاده شده است.

۳-۲-۲- تحلیل همبستگی رتبه‌ای Spearman

ضریب همبستگی رتبه‌ای Spearman صورتی از ضریب همبستگی Pearson است و زمانی به کار برده می‌شود که در صدد بررسی ارتباط آماری بین دو متغیر با داده‌های رتبه‌ای باشیم. در واقع، این ضریب (r_s)، میزان همبستگی میان دو متغیر ترتیبی را نشان می‌دهد. در این ضریب همبستگی به جای استفاده از خود مقادیر متغیرها از رتبه‌های آنان استفاده می‌شود. رابطه‌ی مربوط به ضریب همبستگی رتبه‌ای Spearman به صورت رابطه (۳) تعریف می‌شود.

$$r_s = 1 - \frac{6 \sum_{i=1}^k d_i^2}{n(n^2-1)} \quad (3)$$

در رابطه (۳)، r_s مقدار همبستگی Spearman بین متغیرهای x و y ، n : تعداد آزمودنی‌ها (رتبه‌ها)، k : تعداد دهستان‌ها و d : تفاضل بین رتبه‌ها می‌باشند. در این تحقیق تحلیل همبستگی Spearman به منظور توصیف ارتباط بین بروز بیماری سالک و متغیرهای محیطی و آب و هوایی در بازه زمانی ۰، ۱، ۲، ۳ و ۴ ماه قبل از وقوع بیماری مورد استفاده قرار گرفته شده است.

۳-۳- شبکه عصبی پرسپترون چندلایه^۱

شبکه‌های عصبی مصنوعی یک الگوی ریاضی با هدف بهینه‌سازی و یادگیری بر اساس مفاهیم الهام گرفته از ماهیت مغز و شبکه عصبی انسان می‌باشند (Hagan, Demuth, Beale, & De Jesús, 1996) این شبکه‌ها از تعداد زیادی سلول عصبی به هم پیوسته به عنوان عناصر پردازشی تشکیل شده‌اند که این عناصر به صورت موازی با یکدیگر برای حل یک مسئله

2- Fully-Connected

3- Back-Propagation

1- Multilayer Perceptron Neural Network (MLP)

فصلنامه علمی - پژوهشی اطلاعات جغرافیایی (مهر)
ارائه مدلی برای پیش‌بینی بیماری لیشمانیوز جلدی (سالک) ... / ۱۵

عصبی استفاده شده است (در مرحله آموزش شبکه، به منظور تعیین پارامترهای بهینه مدل، از روش‌های جستجوی شبکه^۱ و اعتبارسنجی متقابل^۲ بر روی ۲۵ درصد از داده‌های آموزشی استفاده شده است. دلیل استفاده از این روش‌ها، تعیین مقادیر بهینه پارامترهای مدل برای کاربرد مورد نظر می‌باشد به گونه‌ای که مدل بتواند داده‌های ارزیابی را با دقت خوبی پیش‌بینی کند. برای کسب اطلاعات بیشتر در رابطه با روش‌های جستجوی شبکه و اعتبارسنجی متقابل، (به *Hsu, Chang, and Lin (2003)* رجوع شود). سپس به منظور ارزیابی مدل آموزش دیده و تهیه نقشه‌های پیش‌بینی و آسیب‌پذیری بیماری سالک، داده‌های مربوط به ۳۰ درصد باقیمانده از کل داده‌ها وارد مدل شدند و دقت شبکه مورد ارزیابی قرار گرفت. در شبکه عصبی پرسپترون چندلایه، مطابق نگاره (۵)، اولین لایه میانی، داده‌های ورودی به همراه وزن مربوطه که بیانگر اهمیت نسبی داده مورد نظر می‌باشد را دریافت می‌کند. همچنین یک بایاس (*b*) به عنوان ارتباط میان وزن‌ها و یک مقدار ثابت غیر صفر به حاصل ضرب هر لایه ورودی با وزن مربوطه‌اش اضافه می‌شود و مقدار به دست آمده در تابع فعال‌سازی قرار گرفته و نتایج حاصل از آن به عنوان خروجی لایه مورد نظر و ورودی لایه بعد در نظر گرفته می‌شود. در نهایت خروجی آخرین لایه میانی به عنوان مقدار لایه خروجی، پاسخ واقعی شبکه را تشکیل می‌دهد (*Hornik et al., 1989*).

۴-۲- بررسی‌های آماری و آمار مکانی

بررسی‌های آماری صورت گرفته در این تحقیق در سال‌های ۱۳۹۴-۱۳۹۲ نشان داد که در طی ۳ سال مورد مطالعه، در مجموع ۲۸۲۹ مورد بیماری سالک در استان گلستان گزارش شده‌اند. از این تعداد موارد بیماری، بیشترین موارد وقوع بیماری متعلق به دهستان‌های گنبدکاووس و مراوه تپه می‌باشند (نگاره ۶).



نگاره ۴: روند کلی تحقیق

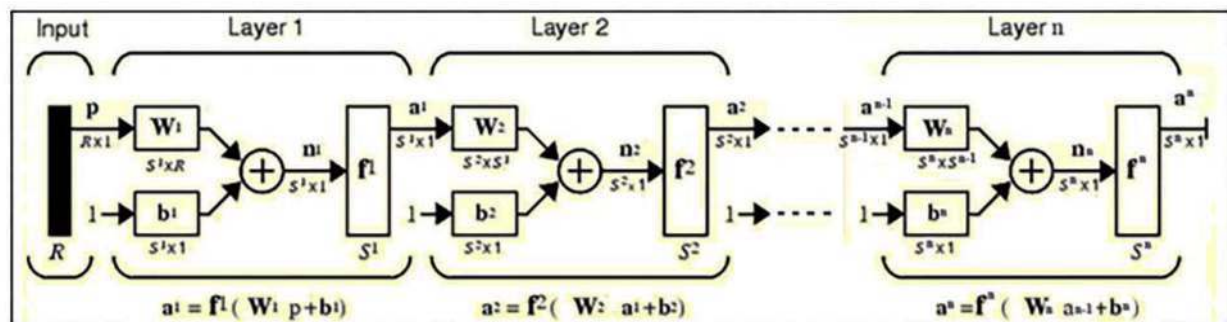
۴-۱- پیاده‌سازی الگوریتم پرسپترون چندلایه جهت مدل‌سازی بیماری سالک

به منظور آموزش شبکه عصبی پرسپترون چندلایه جهت تهیه نقشه پیش‌بینی بیماری سالک بدین صورت عمل شده است که از ۷۰ درصد کل داده‌ها (پارامترهای مؤثر ورودی به همراه نرخ شیوع بیماری سالک) با هدف آموزش شبکه

1- Grid-Search
 2- Cross- Validation

جدول ۲: خصوصیات سنی و جنسیتی بیماران مبتلا به سالک در استان گلستان طی سال‌های ۱۳۹۲-۱۳۹۴

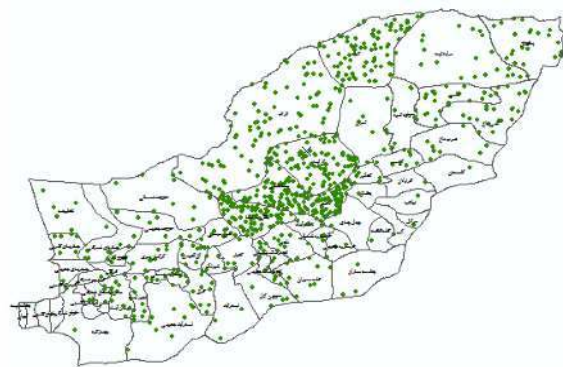
سال	گروه سنی							جنسیت		مجموع
	۰-۴	۵-۹	۱۰-۱۴	۱۵-۲۴	۲۵-۳۴	۳۵-۶۴	+۶۵	مرد	زن	
۱۳۹۲	۳۴۹	۲۷۱	۱۴۶	۴۱۴	۱۸۷	۲۴۲	۱۳	۱۰۱۶	۶۰۶	۱۶۲۲
۱۳۹۳	۱۱۲	۸۶	۴۷	۱۶۱	۱۰۹	۱۲۳	۹	۴۱۳	۲۳۴	۶۴۷
۱۳۹۴	۱۶۲	۸۹	۲۰	۱۰۳	۶۱	۱۱۹	۶	۳۵۴	۲۰۶	۵۶۰



نگاره ۵: شبکه عصبی پرسپترون چندلایه با n لایه میانی (Hornik et al., 1989; Menhaj, 1998)

بیماری می‌شود. خصوصیات سنی و جنسیتی بیماران مبتلا به سالک در جدول (۲) نمایش داده شده است. همچنین بررسی‌های انجام شده نشان داده است که فراوانی بیماری سالک به این صورت بوده است که بروز بیماری از تیر ماه شروع شده و در ماه‌های شهریور، مهر و آبان (فصل پاییز) به بالاترین میزان رسیده است و سپس کاهش یافته است. دلیل چنین نتیجه‌ای را می‌توان شروع فعالیت پشه‌ها به عنوان ناقل بیماری در این ماه‌ها دانست که با نتایج بدست آمده از مطالعه انجام شده توسط چرابین و همکاران (Cherabin, 2012) مطابقت دارد.

طبق جدول (۳)، نتایج حاصل از شاخص Global Moran's I مربوط به هر ۳ سال مورد بررسی روند مشابهی را نشان می‌دهند و بیانگر وجود خودهمبستگی مکانی بین موارد بیماری سالک در استان گلستان می‌باشد. همچنین از نتایج بدست آمده از شاخص Global Moran's I مطابق با جدول (۳) و نگاره (۷) می‌توان به این نکته پی برد که توزیع بیماری در منطقه در هر ۳ سال خوشه‌ای است.



نگاره ۶: توزیع مکانی موارد بیماری سالک در استان گلستان طی سال‌های ۱۳۹۲-۱۳۹۴

با در نظر گرفتن بروز بیماری در مردان و زنان، بررسی‌های آماری نشان دادند که در طی ۳ سال، ۱۷۸۳ مورد بیماری (۶۳ درصد) در مردان و ۱۰۴۶ مورد بیماری (۳۷ درصد) در زنان اتفاق افتاده است. بروز بیشتر بیماری در مردان می‌تواند مربوط به فاکتورهای رفتاری مانند پوشش کمتر و تعاملات بیشتر محیطی و اجتماعی مردان نسبت به زنان باشد که منتج به افزایش تماس با پشه‌های ناقل

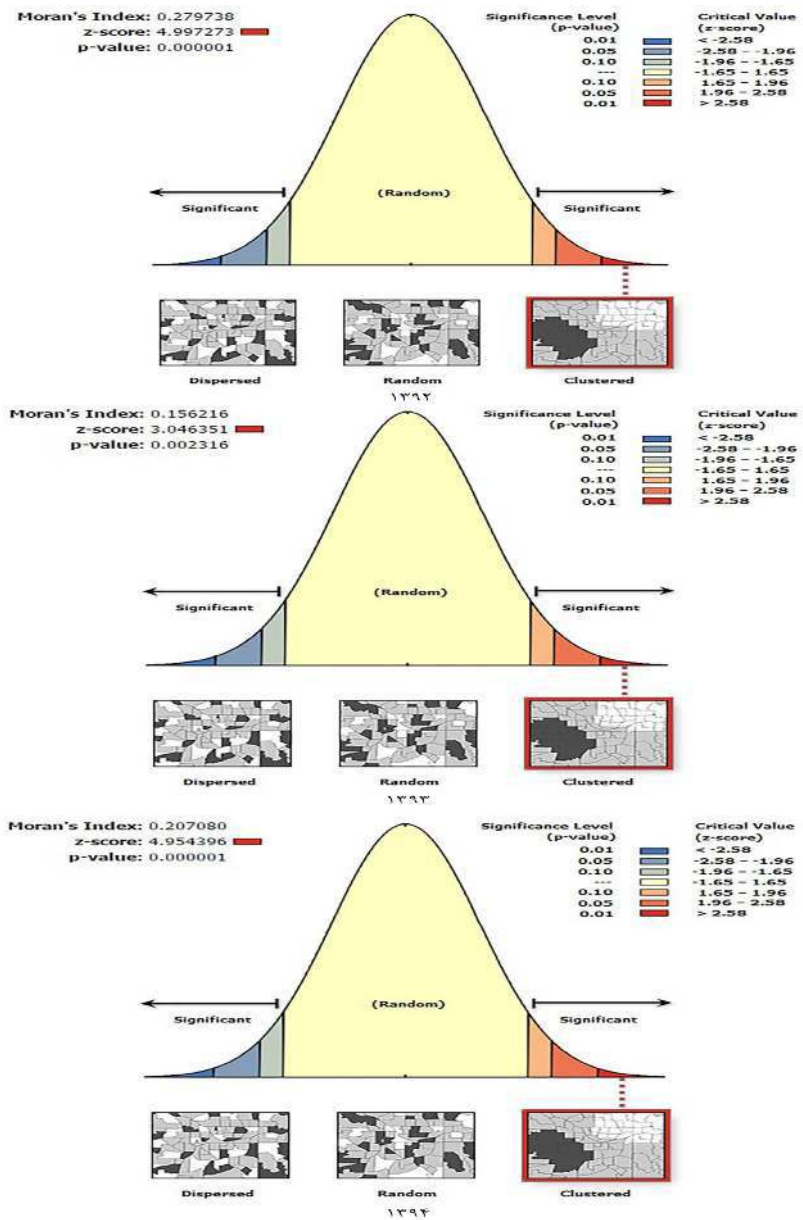
فصلنامه علمی - پژوهشی اطلاعات جغرافیایی (سپهر)
 ارائه مدلی برای پیش‌بینی بیماری لیشمانیوز جلدی (سالک) /... / ۱۷

مشاهده می‌شود میان متغیر NDVI و بروز بیماری سالک در استان گلستان ارتباط معنی‌داری وجود دارد و بروز بیماری سالک در مناطق با NDVI کمتر، بیشتر است. همچنین بین متغیر متوسط رطوبت و بروز سالک ارتباط معنی‌دار وجود داشته و بروز سالک در مناطق با رطوبت کمتر، بیشتر است. نتایج حاصل از تحلیل همبستگی Pearson ارتباط معنی‌داری بین سایر متغیرهای هواشناسی با بروز سالک را نشان نمی‌دهد.

جدول ۳: نتایج شاخص عمومی Moran, s I (۱۳۹۲-۱۳۹۴)

سال	مقدار شاخص	Z- score	P-value	نوع توزیع
۱۳۹۲	۰/۲۸	۵	<۰/۰۱	خوشه‌ای
۱۳۹۳	۰/۱۶	۳/۰۵	<۰/۰۱	خوشه‌ای
۱۳۹۴	۰/۲۱	۴/۹۵	<۰/۰۱	خوشه‌ای

جدول (۴) و (۵) نتایج حاصل از آنالیزهای Pearson و Spearman را نشان می‌دهند. همان‌طور که در جدول (۴)



نگاره ۷: نتایج حاصل از بررسی الگوی توزیع مکانی موارد بیماری سالک با استفاده از شاخص عمومی Moran, s I

جدول ۴: نتایج آنالیز همبستگی Pearson میان متغیرهای محیطی و آب و هوایی با بروز بیماری سالک

متغیرهای مؤثر	پوشش گیاهی	متوسط رطوبت	متوسط دما	مجموع بارش	تعداد روزهای بارانی	تعداد روزهای یخبندان	حداکثر سرعت باد	میزان تبخیر
بروز بیماری سالک	-۰/۵۴۶	-۰/۳۳۴	۰/۱۹۹	-۰/۲۰۴	-۰/۱۷۲	-۰/۱۱۷	۰/۱۰۶	۰/۱۰۸

همبستگی در سطح اطمینان ۰/۰۱ معنی دار است.

جدول ۵: نتایج آنالیز همبستگی Spearman میان متغیرهای محیطی و آب و هوایی با بروز بیماری سالک در فواصل زمانی ۰، ۱، ۲، ۳ و ۴ ماهه

متغیرهای مؤثر	پوشش گیاهی	متوسط رطوبت	متوسط دما	مجموع بارش	تعداد روزهای بارانی	تعداد روزهای یخبندان	حداکثر سرعت باد	میزان تبخیر
بروز بیماری سالک								
همزمان	-۰/۳۹۷	-۰/۴۰۶	۰/۴۹۵	-۰/۱۵۷	-۰/۴۷۲	-۰/۳۱۲	-۰/۴۰۶	۰/۳۲۱
۱ ماه بعد	-۰/۴۴۶	-۰/۷۱۳	۰/۷۴۶	-۰/۳۹۶	-۰/۵۸۳	-۰/۵۵۲	۰/۰۵۴	۰/۵۸۴
۲ ماه بعد	-۰/۵۸۷	-۰/۸۷۴	۰/۷۵۸	-۰/۵۰۲	-۰/۵۹۲	-۰/۶۷۵	۰/۲۲۸	۰/۷۲۵
۳ ماه بعد	-۰/۵۱۱	-۰/۷۸۱	۰/۶۱۴	-۰/۵۳۷	-۰/۴۷۷	-۰/۶۰۲	۰/۳۶۸	۰/۷۰۴
۴ ماه بعد	-۰/۳۲۴	-۰/۴۸۹	۰/۲۷۲	-۰/۳۴۴	۰/۱۵۸	-۰/۴۴۹	۰/۳۳۴	۰/۴۶۳

همبستگی در سطح اطمینان ۰/۰۱ معنی دار است.

۴-۴- ارزیابی نقشه‌ها با استفاده از معیارهای ارزیابی در این تحقیق، به منظور ارزیابی نقشه‌های پیش‌بینی و آسیب‌پذیری نرخ شیوع بیماری سالک حاصل از روش شبکه عصبی پرسپترون چندلایه (بررسی قدرت تشخیص و پیش‌بینی مدل شبکه عصبی پرسپترون چندلایه)، از معیارهای خطای جذر میانگین مربعات (RMSE) و درصد میانگین مطلق خطا (MAPE) و سطح زیر نمودار مشخصه عملکرد سیستم^۳ (AUC) استفاده شده است. معیارهای RMSE و MAPE در علم آمار بیانگر میزان خوب بودن مدل آماری در تخمین متغیر وابسته بر اساس متغیرهای مستقل می‌باشد. مقادیر پایین این معیارها بیانگر این موضوع است که مدل استفاده شده قادر است به طور قابل قبول متغیر وابسته را با در نظر گرفتن متغیرهای مستقل پیش‌بینی کند و مقادیر بالا نشان‌دهنده آن است که نمی‌توان با متغیرهای مستقل در نظر گرفته شده متغیر وابسته را تخمین زد. به عبارت دیگر،

مطابق جدول (۵) نتایج ضریب همبستگی Spearman

میان متغیر NDVI و بروز سالک در فواصل ۰، ۱، ۲، ۳ و ۴ ماهه در سطح استان گلستان، همبستگی قوی و منفی برای متغیر NDVI را نشان می‌دهد. همچنین این ضریب، همبستگی قوی و مثبت میان متغیر متوسط دما و بروز سالک را نشان می‌دهد. در حالی که همبستگی بین متغیر متوسط رطوبت معنی‌دار اما منفی است. از طرفی، متغیر حداکثر سرعت باد ضعیف‌ترین همبستگی مثبت را نشان می‌دهد.

۴-۳- نقشه‌های پیش‌بینی و آسیب‌پذیر تولید شده

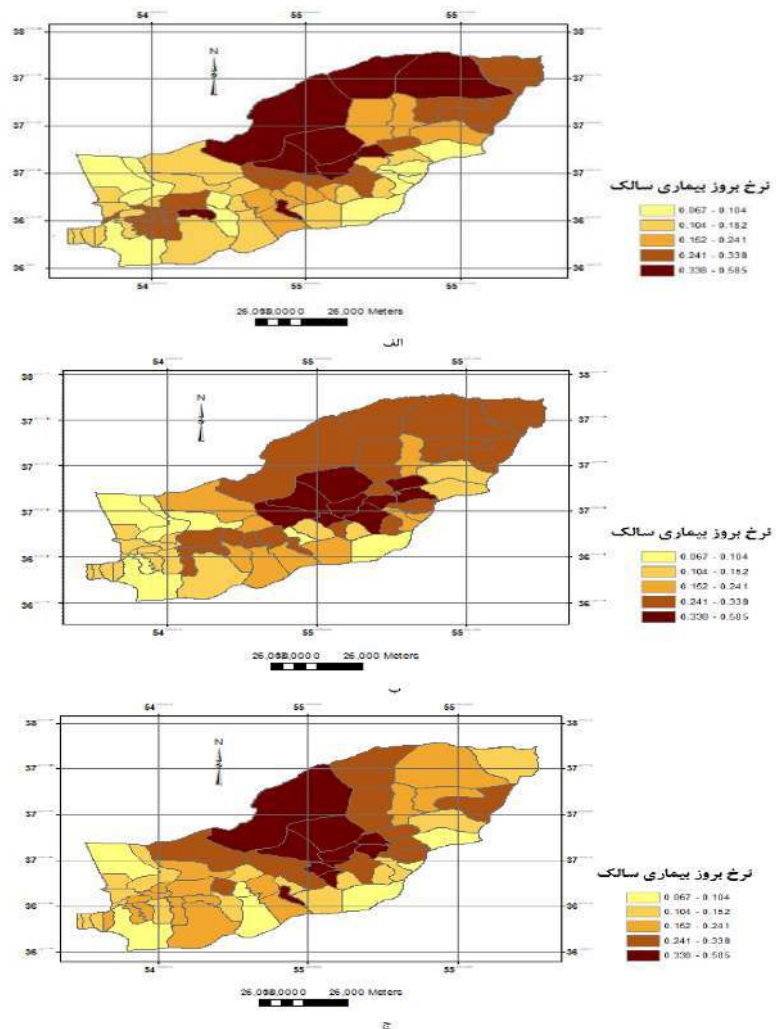
نگاره (۸) نقشه‌های پیش‌بینی و آسیب‌پذیری نرخ بروز بیماری سالک را براساس مدل شبکه عصبی پرسپترون چندلایه نشان می‌دهد. در این نقشه‌ها بیشترین احتمال وقوع بیماری غالباً در قسمت‌های مرکزی و شمال شرقی استان گلستان مشاهده می‌شود. همچنین مطابق با این نقشه‌ها احتمال پراکنش بیماری در تمام سطح استان گلستان وجود دارد.

1- Root Mean Square Error (RMSE)

2- Mean Absolute Percentage Error (MAPE)

3- Receiver Operating Characteristic (ROC)

4- Area Under Curve (AUC)



نگاره ۸: الف) نقشه پیش‌بینی و آسیب‌پذیری نرخ بروز بیماری سالک در سال ۱۳۹۲. ب) نقشه پیش‌بینی و آسیب‌پذیری نرخ بروز بیماری سالک در سال ۱۳۹۳. ج) نقشه پیش‌بینی و آسیب‌پذیری نرخ بروز بیماری سالک در سال ۱۳۹۴

واقعی می‌باشند. معیار ارزیابی AUC نشان دهنده مساحت زیر نمودار منحنی مشخصه عملکرد سیستم می‌باشد که مقداری بین ۰/۵ تا ۱ دارد. هرچه مقدار این عدد بزرگتر و به عدد ۱ نزدیکتر باشد، کارایی الگوریتم پیش‌بینی و دسته‌بندی، مطلوب‌تر ارزیابی می‌شود و مقدار ۰/۵ بیانگر تصادفی بودن نتایج حاصل از الگوریتم می‌باشد. به منظور محاسبه سطح زیر نمودار ROC، مطابق جدول (۶) از دو پارامتر نرخ تشخیص صحیح دسته مثبت^۱ روی محور Y و نرخ تشخیص غلط دسته منفی^۲ روی محور X استفاده می‌شود.

نتایج حاصل از ارزیابی نقشه‌های پیش‌بینی و آسیب‌پذیری

1- True Positive Rate (TPR)

2- False Positive Rate (FPR)

این معیارهای ارزیابی نشان می‌دهند که چگونه تغییرات متغیر وابسته توسط متغیرهای مستقل قابل برآورد می‌باشد. این معیارهای ارزیابی با استفاده از روابط (۴) و (۵) قابل محاسبه می‌باشند (Draper & Smith, 1998; Hyndman & Koehler, 2006; Tofallis, 2015)

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Y_i - y_i)^2}{N}} \quad (4)$$

$$MAPE = \frac{1}{N} * \sum_{i=1}^N \left| \frac{Y_i - y_i}{Y_i} \right| * 100 \quad (5)$$

که در این رابطه: N، تعداد پیش‌بینی‌ها، Y_i ، y_i و Y به

ترتیب بیانگر مقادیر واقعی، پیش‌بینی شده و میانگین مقادیر

جدول ۶: معیار ارزیابی AUC

واقعی	پیش بینی	
	مثبت	منفی
مثبت	صحیح دسته مثبت (TP)	غلط دسته منفی (FN)
منفی	غلط دسته مثبت (FP)	صحیح دسته منفی (TN)
$TPR = \frac{TP}{FN+TP} \quad , \quad FPR = \frac{FP}{TN + FP}$		
TP: بیانگر تعداد رکوردهایی است که دسته واقعی آنها مثبت بوده و الگوریتم دسته بندی نیز دسته آنها را به درستی مثبت تشخیص داده است.		
FP: بیانگر تعداد رکوردهایی است که دسته واقعی آنها منفی بوده و الگوریتم دسته بندی دسته آنها را به اشتباه مثبت تشخیص داده است.		
FN: بیانگر تعداد رکوردهایی است که دسته واقعی آنها مثبت بوده و الگوریتم دسته بندی دسته آنها را به اشتباه منفی تشخیص داده است.		
TN: بیانگر تعداد رکوردهایی است که دسته واقعی آنها منفی بوده و الگوریتم دسته بندی نیز دسته آنها را به درستی منفی تشخیص داده است.		

از پارامترهای محیطی و آب و هوایی مؤثر بر روی مدل آموزش دیده شبکه عصبی پرسپترون چندلایه در پیش بینی نرخ بروز بیماری سالک استفاده شده است.

فرآیند اجرای روش OAT بدین صورت است که مدل به دست آمده به تعداد پارامترهای محیطی و آب و هوایی مورد استفاده در این تحقیق اجرا شده است. با این تفاوت که در هر بار تکرار یک متغیر از بین متغیرهای مؤثر در بروز بیماری از مدل خارج شده و عمل مدل سازی با متغیرهای باقیمانده صورت می گیرد و عملکرد مدل با حذف آن متغیر با استفاده از معیارهای ارزیابی مدل سنجیده می شود (Saltelli et al., 2008). همانطور که در جدول (۸) و نگاره (۱۰) قابل مشاهده است، هنگامی که متغیر پوشش گیاهی از مدل سازی حذف می شود عملکرد مدل را بیش از سایر متغیرها کاهش می دهد. بنابراین در تحقیق حاضر، متغیر پوشش گیاهی مهمترین متغیر در مدل سازی بیماری سالک به شمار می آید. با توجه به نکات فوق، مؤثرترین متغیرها در بروز بیماری سالک بر اساس مدل شبکه عصبی پرسپترون، به ترتیب متغیرهای پوشش گیاهی و متوسط رطوبت هوا می باشند. نتایج بدست آمده از آنالیز حساسیت در این تحقیق با نتایج حاصل از مطالعات انجام شده توسط باویا و همکاران (Bavia et al., 2005)، مظفری و همکاران (Mozaffari et al., 2012) و شیرزادی و همکاران (Shirzadi, Mollalo, & Yaghoobi-Ershadi, 2015) مطابقت دارد.

تولید شده براساس معیارهای مورد نظر در جدول (۷) ارائه شده است. همان طور که از جدول (۷) و نگاره (۹) مشخص است، مدل شبکه عصبی پرسپترون چندلایه قدرت تشخیص بسیار خوبی در پیش بینی وقوع بیماری دارد و حاکی از آن است که مدل به دست آمده در پیش بینی نقاط با ریسک بالای احتمال وقوع بیماری سالک در سطح استان گلستان از کارایی بالا و قابل قبولی برخوردار است.

جدول ۷: ارزیابی نتایج حاصل از مدل شبکه عصبی پرسپترون چندلایه با استفاده از معیارهای ارزیابی

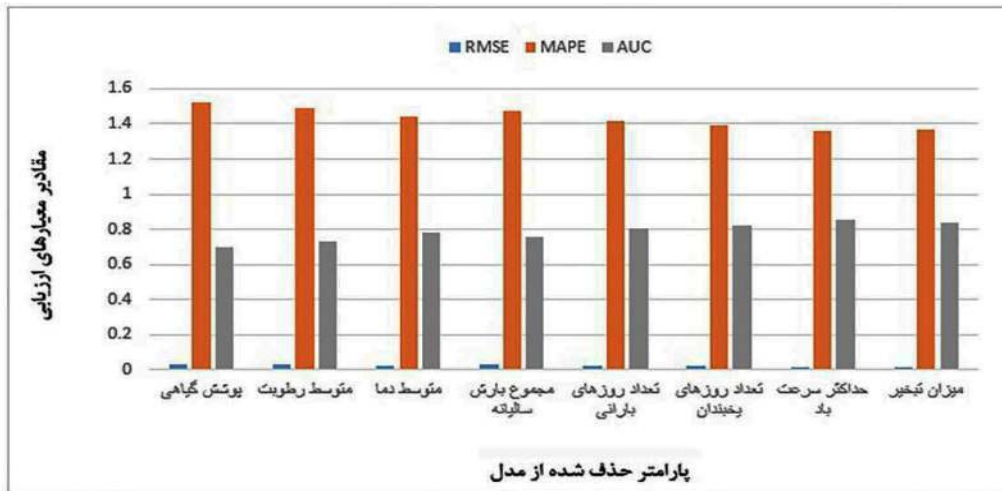
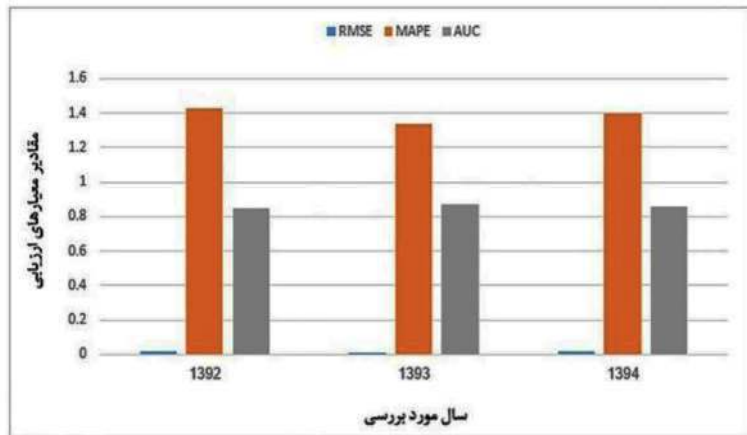
سال	RMSE	MAPE	AUC
۱۳۹۲	۰/۰۱۹	۱/۴۳	۰/۸۴۶
۱۳۹۳	۰/۰۱۳	۱/۳۴	۰/۸۷۳
۱۳۹۴	۰/۰۱۷	۱/۴۰	۰/۸۵۹

۴-۵- آنالیز حساسیت^۱

آنالیز حساسیت مجموعه ای از تکنیک هایی است که تأثیر هر یک از پارامترهای مستقل را بر روی یک متغیر وابسته خاص در یک مدل ریاضی تعیین می کند (Saltelli et al., 2002; Saltelli et al., 2008). روش های بسیاری به منظور اجرای آنالیز حساسیت وجود دارد (Saltelli et al., 2008). در این مطالعه از روش (OAT) one at a time به منظور تعیین تأثیر هر یک

فصلنامه علمی - پژوهشی اطلاعات جغرافیایی (سیر)
 ارائه مدلی برای پیش‌بینی بیماری لیشمانیوز جلدی (سالک) ... / ۲۱

نگاره ۹: ارزیابی نتایج حاصل از مدل با استفاده از معیارهای ارزیابی



نگاره ۱۰: نتایج حاصل از آنالیز حساسیت OAT

وجود خودهمبستگی مکانی میان موارد بروز بیماری سالک بوده و نشان داد که توزیع بیماری سالک در منطقه برای هر سال مورد مطالعه خوشه‌ای می‌باشد. همچنین نتایج حاصل تحلیل همبستگی پیرسون و رتبه‌ای اسپیرمن گویای این حقیقت بودند که عوامل محیطی و آب و هوایی در وقوع بیماری سالک نقش بسیار مهمی را ایفا می‌کنند. به منظور تهیه نقشه‌های پیش‌بینی و آسیب‌پذیری بیماری سالک در سطح دهستان‌های استان گلستان از متغیرهای محیطی و آب و هوایی شامل پوشش گیاهی، متوسط رطوبت هوا، متوسط دمای هوا، مجموع بارش سالیانه، تعداد روزهای بارانی، تعداد روزهای یخبندان، حداکثر سرعت باد و میزان تبخیر استفاده شده است. در این مدل‌سازی روش شبکه عصبی پرسپترون چندلایه به کار برده شده است. نقشه‌های

جدول ۸: نتایج حاصل از آنالیز حساسیت OAT

متغیر خارج شده از مدل	مدل شبکه عصبی پرسپترون چندلایه		
	RMSE	MAPE	AUC
پوشش گیاهی	۰/۰۳۳	۱/۵۲	۰/۷۰۲
متوسط رطوبت	۰/۰۳۱	۱/۴۹	۰/۷۳۵
متوسط دما	۰/۰۲۵	۱/۴۴	۰/۷۷۹
مجموع بارش سالیانه	۰/۰۲۸	۱/۴۷	۰/۷۵۶
تعداد روزهای بارانی	۰/۰۲۳	۱/۴۲	۰/۸۰۴
تعداد روزهای یخبندان	۰/۰۱۹	۱/۳۹	۰/۸۱۸
حداکثر سرعت باد	۰/۰۱۵	۱/۳۶	۰/۸۵۴
میزان تبخیر	۰/۰۱۷	۱/۳۷	۰/۸۴۱

۵- نتیجه‌گیری و ارائه پیشنهادات

در این تحقیق، تحلیل مکان-آماري استفاده شده بیانگر

مدت (۳ سال) به دلیل فقدان و یا کامل نبودن داده‌های بیماری مربوط به سال‌های قبل بوده است، توصیه می‌شود که از داده‌های بیماری سالک برای دوره زمانی بلند مدت به منظور افزایش دقت مدل‌سازی استفاده شود.

منابع و مآخذ

- 1- AhangarCani, Farnaghi & Shirzadi, M., M., M.R. (2016). Predictive Map of Spatio-Temporal Distribution of Leptospirosis Using Geographical Weighted Regression and Multilayer Perceptron Neural Network Methods. *Journal of Geomatics Science and Technology*, 6(2), 79-98.
- 2- Antonialli, Torres, Paranhos Filho & Tolezano, S. A. C., T. G., A. C., J. E. (2007). Spatial analysis of American visceral leishmaniasis in Mato Grosso do Sul state, Central Brazil. *Journal of infection*, 54(5), 509-514.
- 3- Assimina, Charilaos & Fotoula, Z., K., B. (2008). Leishmaniasis: an overlooked public health concern. *Health Science Journal*, 2(4).
- 4- Bavia, Carneiro, da Costa Gurgel, Filho & Barbosa, M., D., H., C. M., M. R. (2005). Remote sensing and geographic information systems and risk of American visceral leishmaniasis in Bahia, Brazil. *Parassitologia*, 47(1), 165.
- 5- Bayatani & Sadeghi, A., A. (2012). Spatial Analysis of Environmental Factors of Cutaneous Leishmaniasis in Iran Using GIS. *Hakim Research Journal*, 15, 158-165.
- 6- Chaves & Pascual, L. F., M. (2006). Climate cycles and forecasts of cutaneous leishmaniasis, a nonstationary vector-borne disease. *PLoS Med*, 3(8), e295.
- 7- Cherabin, Palideh, Gharavi, Gharavi & mahmood, m. A. S., A, A, A. (2012). Epidemiological characteristics of Cutaneous Leishmaniasis in Maraveh tapeh district, Golestan province during 2006-2010. *Journal of zabol university of medical sciences and health services*, 4(1), 19-27.
- 8- Draper & Smith, N., H. (1998). *Applied regression analysis*, Wiley Interscience. New York, 505-553.
- 9- Gordis, L. (2000). *Epidemiology WB Saunders*. Philadelphia, PA.
- 10- Hagan, Demuth, Beale & De Jesús, M. T., H. B., M.

پیش‌بینی و آسیب‌پذیری تولید شده بیانگر این نکته بودند که مدل پرسپترون چندلایه قادر بوده است تا با دقت قابل قبولی نرخ بروز بیماری سالک را پیش‌بینی نماید.

RMSE1392 = 0.019, RMSE1393 = 0.013, RMSE1394 = 0.017, MAPE1392 = 1.43, MAPE1393 = 1.34, MAPE1394 = 1.40, AUC1392 = 0.846, AUC1393 = 0.873, AUC1394 = 0.859

همچنین نتایج تحقیق براساس معیارهای ارزیابی نشان دادند که از میان متغیرهای مؤثر بر بروز بیماری، متغیرهای پوشش گیاهی و متوسط رطوبت هوا بیشترین تأثیر را در بروز بیماری سالک داشته‌اند (جدول ۸ و نگاره ۱۰). نقشه‌های پیش‌بینی و آسیب‌پذیری تهیه شده در این تحقیق می‌توانند برای مشخص نمودن نحوه توزیع بیماری سالک در استان گلستان مورد استفاده مدیران و تصمیم‌گیران قرار گیرند. در ضمن برنامه‌ریزی دقیق و تخصیص بودجه و امکانات به مناطقی که بیشتر در معرض ابتلا به این بیماری هستند، از اقدامات لازم در زمینه پیشگیری و کنترل این بیماری می‌باشد. به عنوان کارهای آینده پیشنهاد می‌شود که از سایر متغیرهای محیطی و اجتماعی - اقتصادی از جمله توپوگرافی، تراکم جمعیت، مهاجرت از مناطق غیر بومی به مناطق بومی و بالعکس، تغییرات ساخت بشر همچون کویرزدایی، سدسازی و غیره به منظور بهبود در فرآیند مدل‌سازی بیماری سالک استفاده شود. همچنین توصیه می‌شود که چنین مطالعه‌ای در سایر استان‌های بومی و در معرض خطر بالای بیماری سالک مورد استفاده قرار گیرد. با توجه به اینکه مطالعه حاضر، به دلیل فراهم نبودن داده‌های کامل ماهیانه بیماری سالک، به صورت سالیانه انجام شده است، پیشنهاد می‌شود از آنجایی که بیماری‌های واگیر از جمله بیماری سالک تحت تأثیر تغییرات عوامل محیطی و اقلیمی، از زمانی به زمان دیگر بر میزان بروز و تشدید آنها مؤثر می‌باشند، مطالعه حاضر به صورت فصلی و یا ماهانه صورت گیرد. همچنین با توجه به اینکه یکی دیگر از محدودیت‌های این تحقیق استفاده از داده‌های بیماری کوتاه

- 22- Mohammady& Delavar, S., M. (2014). Urban Expansion Modeling with Logistic Regression. *Journal of Geomatics Science and Technology*, 4(2), 77-86.
- 23- MOLLALO, A. (2014). Spatio-Temporal Analyses and Modeling of Cutaneous Leishmaniasis Disease, Alimohammadi., Malek., A., M.R., K.N. Toosi University of Technology, Civil-Surveying Engineering In GIS.
- 24- Mollalo, Alimohammadi, Shirzadi& Malek, A., A., M., M. (2015). Geographic Information System-Based Analysis of the Spatial and Spatio-Temporal Distribution of Zoonotic Cutaneous Leishmaniasis in Golestan Province, North-East of Iran. *Zoonoses and public health*, 62(1), 18-28.
- 25- Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37(1/2), 17-23.
- 26- Morrone, Pitidis, Pajno, Dassoni, Latini, Ab Barnabas& Padovese, A., A., M. C., F., O., G., V. (2011). Epidemiological and geographical aspects of leishmaniasis in Tigray, northern Ethiopia: a retrospective analysis of medical records, 2005-2008. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 105(5), 273-280.
- 27- Mozaffari, BAKHSHIZADEH& GHEIBI, G., F., M. (2012). Analysis relationship between vegetation cover and Salak skin disease in Yazd-Ardakanplain. 22(4), 167-178.
- 28- Organization, W. H. (2009). A human rightsbased approach to neglected tropical diseases: WHO.
- 29- Rajabi, Mansourian, Pilesjö& Bazmani, M., A., P., A. (2014). Environmental modelling of visceral leishmaniasis by susceptibility-mapping using neural networks: a case study in north-western Iran. *Geospatial health*, 9(1), 179-191.
- 30- Rajabi, Pilesjö, Shirzadi, Fadaei& Mansourian, M., P., M. R., R., A. (2016). A spatially explicit agent-based modeling approach for the spread of CutaneousLeishmaniasis disease in central Iran, Isfahan. *Environmental Modelling & Software*, 82, 330-346.
- 31- Ramezankhani, Hosseini, Sajjadi, Khoshabi& Ramezankhani, R., A., N., M., A. (2017). Environmental risk factors for the incidence of cutaneous leishmaniasis in an endemic area of Iran: A GIS-based approach. Spatiol. H., O. (1996). *Neural network design*, Vol. 20, PWS publishing company Boston.
- 11- Hassan, Kenawy, Kamal, Abdel Sattar& Sowilem, A., M., H., A., M. (2003). GIS-based prediction of malaria risk in Egypt. *Eastern Mediterranean Health Journal*, 9(4), 11.
- 12- Herbreteau, Demoraes, Khaungaew, Hugot, Gonzalez, Kittayapong& Souris, V., F., W., J.-P., J.-P., P., M. (2006). Use of geographic information system and remote sensing for assessing environment influence on leptospirosis incidence, Phrae province, Thailand. *International Journal of Geoinformatics*, 2(4), 43-50.
- 13- Holakouie-Naieni, Mostafavi, Bolorani, Mohebalil& Pakzad, K., E., A. D., M., R. (2017). Spatial modeling of cutaneous leishmaniasis in Iran from 1983 to 2013. *Acta Tropica*, 166, 67-73.
- 14- Hornik, Stinchcombe& White, K., M., H. (1989). Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5), 359-366.
- 15- Hsu, Chang& Lin, C.-W., C.-C., C.-J. (2003). A practical guide to SVM classification. Department of Computer Science and Information Technology, National Taiwan University. Paper available at <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>, 16).
- 16- Hyndman & Koehler, R. J., A. B. (2006). Another look at measures of forecast accuracy. *International journal of forecasting*, 22(4), 679-688.
- 17- Masoomy & Mesgari, Z., M. S. (2006). PREDICTION OF SKIN CANCER EPIDEMIOLOGY FOR DECISION MAKING USING GEOSTATISTICAL ANALYSES. Paper presented at the ASPRS 2006 Annual Conference, Reno, Nevada.
- 18- McLafferty, S. L. (2003). GIS and health care, *Annual review of public health*, 24(1), 25-42.
- 19- Menhaj, M. (1998). Fundamentals of neural networks. *Computational intelligence*, 1(1).
- 20- Mesgari & Masoomi, M., Z. (2008). GIS applications in public health as a decision making support system and its limitation in Iran. *World appl Sci J (Supple 1)*, 3, 73-77.
- 21- Mitchell, A. (2005). *The ESRI guide to GIS analysis, Volume 2, Spatial Measurements and Statistics*. Redlands: CA: Esri Press.

1352-1362.

44- Williams & Hinton, D., G. (1986). Learning representations by back-propagating errors. *Nature*, 323, 533-536.

45- Yilmaz, I. (2009). Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: a case study from Kat landslides (Tokat—Turkey). *Computers & Geosciences*, 35(6), 1125-1138.

46- Zare, Shamszadeh & Najjari, M., P., A. (2006). Providing the opportunity to use GIS in decision-making in the health sector management. *Hakim*, 9(1), 58-63.

tial and Spatio-temporal Epidemiology, 21, 57-66.

32- Raoufy, Vahdani, Alavian, Fekri, Eftekhari, & Gharibzadeh, M. R., P., S. M., S., P., S. (2011). A novel method for diagnosing cirrhosis in patients with chronic hepatitis B: artificial neural network approach. *Journal of medical systems*, 35(1), 121-126.

33- Rassi, Javadian, Jalali, Motazedian & Vatndost, Y., E., M., M., H. (2004). Investigation on Zoonotic Cutaneous Leishmaniasis, Southern Iran. *Iranian Journal of Public Health*, 33(1), 31-35.

34- Ribeiro Jr, & Diggle, P., P. (2001). *geoR: A package for geostatistical analysis*, 2001: ISSN.

35- Ripley, B. (1981). *Spatial statistics*, 252 pp: John Wiley, New York.

36- Ruiz, Tedesco, McTighe, Austin & Kitron, M. O., C., T. J., C., U. (2004). Environmental and social determinants of human risk during a West Nile virus outbreak in the greater Chicago area, 2002. *International Journal of Health Geographics*, 3(1), 8.

37- Saltelli, A. (2002). Sensitivity analysis for importance assessment. *Risk Analysis*, 22(3), 579-590.

38- Saltelli, Ratto, Andres, Campolongo, Cariboni, Gatelli & Tarantola, A., M., T., F., J., D., S. (2008). *Global sensitivity analysis: the primer*: John Wiley & Sons.

39- Shirzadi, Mollalo & Yaghoobi-Ershadi, M. R., A., M. R. (2015). Dynamic relations between incidence of zoonotic cutaneous leishmaniasis and climatic factors in Golestan Province, Iran. *Journal of arthropod-borne diseases*, 9(2), 148.

40- Skapura, D. M. (1996). *Building neural networks*: Addison-Wesley Professional.

41- Soccol, de Castro, e Schühli, de Carvalho, Marques, de Fátima Pereira & Membrive, V. T., E. A., G. S., Y., E., E., N. (2009). A new focus of cutaneous leishmaniasis in the central area of Paraná State, southern Brazil. *Acta Tropica*. 111(3), 308-315.

42- Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46(sup1), 234-240.

43- Tofallis, C. (2015). A better measure of relative prediction accuracy for model selection and model estimation. *Journal of the Operational Research Society*, 66(8),