



مطالعه تطبیقی دقت مدل‌های زمین‌آماری و رگرسیونی جهت پهنه‌بندی دما در مرکز و شمال ایران

زهرا آقاشریعتمداری^{۱*}، مهسا اتحادی^۲، خلیل قربانی^۳

تاریخ دریافت: ۱۳۹۶/۱۲/۱۹

تاریخ پذیرش: ۱۳۹۷/۱۱/۲۱

چکیده

دمای هوا از مهم‌ترین متغیرهای مورد نیاز در مطالعات اقلیم‌شناسی کشاورزی است. داده‌های دیدبانی شده دما ماهیت نقطه‌ای دارند و جهت برآورد منطقه‌ای این کمیت نیاز به کاربرد روش‌های درون‌یابی است که تغییرات دما بر حسب ارتفاع و عرض جغرافیایی را برآورد کنند. در این پژوهش، کارایی روش‌های کریجینگ، کوکریجینگ، رگرسیون وزن‌دار جغرافیایی و رگرسیون چندمتغیره خطی برای پهنه‌بندی میانگین ماهانه دما ۵۶ ایستگاه همدیدی واقع در مرکز و شمال ایران، مورد ارزیابی قرار گرفته است. تحلیل آماری نتایج نشان داد که رگرسیون وزن‌دار جغرافیایی در ماه دسامبر بیش‌ترین اختلاف را با دیگر روش‌های مورد مطالعه دارد به طوری که ریشه میانگین مربعات خطا در این روش برابر ۰/۸۳ درجه سانتی‌گراد می‌باشد، در حالی که این مقدار در روش کوکریجینگ که در رتبه بعدی قرار دارد برابر با ۲/۳۸ درجه سانتی‌گراد است. کم‌ترین میانگین مقادیر RMSE تمام ماه‌های سال، متعلق به روش رگرسیون وزن‌دار جغرافیایی به میزان ۱/۲ درجه سانتی‌گراد و پس از آن به ترتیب روش‌های رگرسیون چند متغیره خطی (۲/۲۴ درجه سانتی‌گراد)، کریجینگ (۲/۵۲ درجه سانتی‌گراد) و کوکریجینگ (۲/۸۶ درجه سانتی‌گراد) در اولویت‌های بعدی قرار دارند، لذا روش وزن‌دار جغرافیایی برای پهنه‌بندی دما در این منطقه، مناسب‌تر می‌باشد. این رهیافت، به دلیل در نظر گرفتن گرادین‌های محلی ارتفاعی، مقادیر دما در مناطق مرتفع فاقد ایستگاه هواشناسی را دقیق‌تر پیش‌بینی می‌کند.

واژه‌های کلیدی: درون‌یابی موضعی، دما، رگرسیون چندمتغیره، رگرسیون وزن‌دار جغرافیایی

مقدمه

اندازه‌گیری شده در ایستگاه‌ها، برای تحلیل‌های منطقه‌ای، به منطقه مطالعاتی تعمیم داده شوند و توزیع آن‌ها در پهنه جغرافیایی ترسیم شود. روش‌های مختلفی به این منظور توسط محققین مختلف مورد بررسی قرار گرفته است که از آن جمله می‌توان به روش‌های درون‌یابی اشاره کرد. روش‌های درون‌یابی با استناد به این قانون جغرافیایی که هر پدیده با پدیده دیگری مرتبط است، اما پدیده‌های نزدیک نسبت به پدیده‌های دور به هم وابسته‌تر هستند، از یک سری نقاط با مقادیر معلوم در اطراف نقاط با مقادیر مجهول استفاده کرده و مقادیر آن‌ها را تخمین می‌زنند (Tobler, 1970). چگونگی مشارکت نقاط معلوم در تخمین مقادیر مجهول سبب می‌شود تا روش‌های درون‌یابی به دو دسته سراسری و موضعی تقسیم‌بندی شوند. در درون‌یابی سراسری، برای برآورد نقاط مجهول از داده‌های تمام نقاط اندازه‌گیری شده استفاده می‌شود، اما در مدل موضعی، از داده‌های همه نقاط برداشت شده برای برآورد مقدار نقطه مجهول استفاده نمی‌شود. در این روش

دما و تغییرات آن، در مکان‌های مختلف و در طول زمان، اساسی‌ترین عامل تغییرات آب و هوایی است و می‌تواند به عنوان پایه‌ای برای درک فرآیندهایی همچون تبخیرتعرق و توزیع پوشش گیاهی در نظر گرفته شود. آگاهی از توزیع مکانی این متغیر به منظور بررسی روند تغییر اقلیم و اثر گلخانه‌ای مورد نیاز است. متغیرهای هواشناسی از جمله دمای هوا به دلیل هزینه‌بر بودن نمونه‌برداری از ایستگاه‌های هواشناسی که به خصوص در نواحی کوهستانی به طور گسسته و نامنظم توزیع شده‌اند، جمع‌آوری می‌شوند. ضرورت ایجاب می‌کند که اطلاعات

^۱ استادیار گروه آبیاری و آبادانی، پردیس کشاورزی و منابع طبیعی دانشگاه تهران

(*نویسنده مسئول: zagha@ut.ac.ir)

DOI: 10.22125/agmj.2019.148887.1034

^۲ دانش‌آموخته کارشناسی ارشد هواشناسی کشاورزی، دانشگاه تهران

^۳ دانشیار گروه مهندسی آب، دانشکده کشاورزی، دانشگاه کشاورزی و منابع طبیعی گرگان

دقت برآورد مفید واقع شود. امکان دستیابی به این روش‌ها در نرم‌افزارهای سیستم‌های اطلاعات جغرافیایی، راه را برای کاربری و ارزیابی این روش‌ها هموار ساخته است (Mehdizadeh, 2002). با توجه به مطرح شدن روش‌های زمین‌آمار محققین مختلفی در سراسر جهان به استفاده از این شیوه در درون‌یابی متغیرهای هواشناسی روی آوردند. به عنوان مثال Hevesi et al., (1992) برای برآورد و درون‌یابی بارندگی متوسط سالانه در منطقه کوه‌های یوکا واقع در جنوب نوادا و جنوب شرقی کالیفرنیا روش‌های عکس فاصله، عکس مجذور فاصله، عکس مکعب فاصله، کریجینگ، همبستگی خطی، همبستگی خطی لگاریتمی، متوسط همسایگی و کوکریجینگ را بررسی نمودند و روش‌هایی که از ارتفاع به عنوان متغیر کمکی استفاده می‌کنند، و در راس آن‌ها کوکریجینگ را روش برتر معرفی کردند. (Abtew et al., 1993) نیز روش‌های زمین‌آمار را برای آنالیز مکانی بارندگی ۲۵ ایستگاه با دوره‌های آماری متفاوت، از سال ۱۹۲۹ الی ۱۹۹۲ ماهانه با حداقل طول دوره آماری ۱۵ سال در جنوب فلوریدا بررسی کردند. در این راستا روش‌های تیسن، روش وزنی عکس فاصله، چندجمله‌ای محلی درجه دو، کریجینگ عام و کریجینگ معمولی را مورد آزمون قرار دادند و روش‌های چندجمله‌ای محلی و کریجینگ معمولی را جزء مناسب‌ترین روش‌های برآورد بارندگی ماهانه معرفی کردند. مقایسه دیگری نیز بین روش‌های کوکریجینگ، کریجینگ معمولی، کریجینگ عام، کریجینگ معمولی، نزدیک‌ترین همسایگی، عکس مجذور فاصله و روش گرادیان همراه با عکس مجذور فاصله (GIDS¹) توسط (Nalder and Wein (1998) به منظور بررسی توزیع مکانی بارندگی و دمای ماهانه در جنگل‌های شمال کانادا انجام شد. آن‌ها در این مطالعه، از متوسط دما و بارندگی ماهانه در ۳۲ ایستگاه با دوره آماری سی ساله استفاده کردند. نتایج به دست آمده نشان می‌دهد که روش GIDS برای بارندگی ماهانه از سایر روش‌ها مناسب‌تر است و روش‌های کریجینگ رونددار، نزدیک‌ترین همسایگی، کوکریجینگ، عکس مجذور فاصله، کریجینگ عام و کریجینگ معمولی، برای دمای ماهانه، و روش‌های کوکریجینگ، عکس مجذور فاصله، نزدیک‌ترین همسایگی، کریجینگ معمولی برای درون‌یابی بارندگی ماهانه به ترتیب در مراتب بعدی قرار دارند. Benavides et

ممکن است، تعداد نقاط محدودی که تا نقطه مورد نظر کمترین فاصله را دارند، در درون‌یابی به کار گرفته شوند، و یا از داده‌هایی استفاده شود که تا نقطه مورد نظر کم‌ترین فاصله را دارند. به بیانی دیگر در این روش نقاط همسایه بر اساس تعداد یا فاصله تعریف می‌شوند. اگر تغییرات مکانی متغیر زیاد باشد، مدل‌های موضعی نتایج بهتری به دست می‌دهد. با این حال تعداد نقاط همسایه که برای برآورد در یک نقطه معین به کار گرفته می‌شوند، در کیفیت نتایج تأثیر به‌سزایی دارد (Ghorbani, 2013). برخی روش‌های درون‌یابی با برقراری روابط آمار کلاسیک بین مقادیر نقاط معلوم، اقدام به تخمین مقادیر مجهول می‌کنند. آن‌ها را درون‌یاب‌های قطعی می‌نامند. اما برخی از روش‌ها با استفاده از ویژگی‌های وابستگی مکانی متغیر مورد نظر اقدام به تخمین مقادیر نقاط مجهول می‌کنند. آن‌ها را درون‌یاب‌های زمین‌آماری می‌نامند (Ghorbani, 2013). با گسترش کاربرد رایانه، استفاده از روش‌های پیچیده‌ای همچون روش‌های زمین‌آمار فراهم شده است. با استفاده از این روش‌ها، برآورد متغیر مورد نظر در نقاط فاقد آمار امکان‌پذیر می‌باشد. روش‌های فوق‌الذکر با متغیرهای پیوسته سروکار دارند و برای تخمین متغیرهای تصادفی پیوسته استفاده می‌شوند. کاربرد این روش‌ها ابتدا در علوم معدن شروع شد و سپس در سایر علوم، از جمله ژئوشیمی، زمین‌شناسی، ژئوفیزیک، مکانیک سنگ، تخمین ذخایر هیدروکربن، هیدرولوژی و منابع آب، توسعه پیدا نمود (Mehdizadeh, 2002). به دنبال روند تکاملی روش‌های زمین‌آماری مورد استفاده در تخمین ذخایر معدنی و پژوهش‌های افرادی چون کریگ و سیشل، پژوهشگری فرانسوی به نام ماترون، با انتشار مقاله‌ای در سال ۱۹۶۲، پایه‌های زمین‌آمار را بنا کرد (Rahimi, 2000). این روش‌ها به طور موفقیت‌آمیز، برای شماری از متغیرهایی که ساختار مکانی دارند، از جمله متغیرهای هیدرولوژیکی به کار می‌روند و هر جا که به حل مسائل میان‌یابی، میانگین‌گیری و طراحی بهینه شبکه‌های نمونه‌برداری نیاز باشد، می‌توان از این روش‌ها کمک گرفت. در واقع مطابق با فرضیات زمین‌آماری، احتمال میزان تشابه بین مقادیر مربوط به نمونه‌های نزدیک‌تر بیشتر است. زمین‌آمار با در نظر گرفتن همبستگی و ساختار مکانی داده‌ها و قابلیت استفاده از روابط بین متغیرها در حل مسائل، می‌تواند در افزایش

¹ Gradient Plus Inverse Distance Squared

روش کریجینگ که در رتبه دوم قرار دارد، اختلاف زیادی دارد. بر این اساس مناسب‌ترین روش برای ترسیم هم‌بارش در استان گیلان روش رگرسیون وزن‌دار جغرافیایی معرفی شد. همچنین Ghorbani and Aghashariatmadari (2014) با در نظر گرفتن میانگین ۳۰ ساله دمای هوا و رطوبت نسبی ۲۴۰ ایستگاه همدیدی و اقلیم‌شناسی در ایران، دقت روش رگرسیون وزن‌دار جغرافیایی با سایر روش‌های درون‌یابی در پهنه‌بندی این دو پارامتر اقلیمی را مورد ارزیابی قرار دادند. نتایج این بررسی در مورد دمای هوا که تحت تاثیر ارتفاع می‌باشد، نشان داد که نتایج درون‌یابی به کمک روش رگرسیون وزن‌دار جغرافیایی با دیگر روش‌های درون‌یابی اختلاف معنی‌داری دارد، در حالی‌که این اختلاف در مورد رطوبت نسبی مشاهده نشد و این می‌تواند به دلیل نبود رابطه مشخصی بین ارتفاع و رطوبت نسبی باشد. (Georganos et al., 2017) از روش رگرسیون وزن‌دار جغرافیایی برای بررسی رابطه بارندگی- پوشش گیاهی نرمال شده در مناطق نیمه‌خشک ساحل استفاده کردند و نشان دادند که روش رگرسیون وزن‌دار جغرافیایی نسبت به روش رگرسیون بر مبنای حداقل مربعات رابطه بارندگی و پوشش گیاهی را به صورت معنی‌داری بهتر توجیه می‌کند. (Zhai et al., 2018) روش‌های مدل‌سازی بر اساس رگرسیون وزن‌دار جغرافیایی، رگرسیون وزن‌دار جغرافیایی در ترکیب با تحلیل مولفه‌های اصلی را به منظور پهنه‌بندی تغییرات مقدار ذرات آلاینده PM2.5 مورد مقایسه قرار دادند و به این نتیجه رسیدند که روش ترکیبی مورد استفاده نتیجه قابل اعتمادتری را حاصل می‌سازد. (Zhou et al., 2018) از روش رگرسیون وزن‌دار جغرافیایی به منظور ارزیابی اثرات عوامل مختلف موثر بر آلودگی هوا با هدف جلوگیری و اتخاذ سیاست‌های محیطی استفاده کردند. Johnson et al., (2019) رابطه بین مخاطرات اقلیمی نظیر خطر سیل و امکان دسترسی مکانی به سرمایه‌گذاری‌های کوچک را با استفاده از روش‌های رگرسیون معمولی و رگرسیون وزن‌دار جغرافیایی مورد بررسی قرار دادند و نشان دادند که رگرسیون وزن‌دار جغرافیایی هدف مورد نظر را با دقت بالاتری برآورد می‌کند. بنابراین وجود روابط مکانی بین متغیرهای اقلیمی با عوارض طبیعی مانند ارتفاع باعث شده تا جستجوی گرادیان‌ها اولین گزینه در پهنه‌بندی آن‌ها باشد. اما در مناطق بزرگ وجود گرادیان‌های متعدد

al., (2006) مدل‌های برآورد دمای هوا، شامل پنج مدل زمین‌آماری و دو مدل رگرسیونی را در منطقه آستوریاس در شمال اسپانیا مقایسه کرده و به این نتیجه رسیدند که، روش رگرسیون چند متغیره که عوامل ارتفاع، عرض جغرافیایی، فاصله از دریا و تابش ورودی خورشید به سطح زمین را در نظر می‌گیرد، برآورد بهتری از دمای هوا در سطح زمین دارد. استفاده از روش‌های رگرسیونی نیز در پهنه‌بندی متغیرهای هواشناسی متداول بوده است. می‌توان میزان متغیر وابسته را در روش‌های رگرسیونی در صورت برقراری رابطه بین یک متغیر وابسته با یک یا چند متغیر مستقل در کل پهنه تخمین زد. در الگوهای رگرسیونی نیز استفاده از روش‌های محلی باعث شده تا جزئیاتی که در روش‌های سراسری نادیده گرفته می‌شوند، تا حدی به تصویر کشیده شود. روش رگرسیون وزن‌دار جغرافیایی یکی از روش‌های رگرسیونی محلی است. (Bostan and Akyurek 2007) برای آشکارسازی توزیع مکانی میانگین سالانه بارش و دمای هوا از روش‌های درون‌یابی کوکریجینگ و رگرسیون وزن‌دار جغرافیایی استفاده کردند و با مقایسه همبستگی مقادیر پیش‌بینی و اندازه‌گیری‌شده به این نتیجه رسیدند که روش رگرسیون وزن‌دار جغرافیایی برای دما و بارش نسبت به روش کوکریجینگ درون‌یابی با دقت بالاتری را ارائه می‌کند. همچنین Gundogdu and Esen (2010) سه روش کریجینگ، کوکریجینگ و رگرسیون وزن‌دار جغرافیایی را برای پهنه‌بندی میانگین ۲۵ ساله بارش سالانه در ترکیه بررسی کردند. آن‌ها با محاسبه ضریب همبستگی بین مقدار پیش‌بینی شده و الگو شده نتیجه گرفتند، که روش رگرسیون وزن‌دار جغرافیایی کم‌ترین مقدار خطای پیش‌بینی را دارا می‌باشد. در ایران نیز (Ghorbani 2013) تغییرات مکانی بارش سالانه در استان گیلان را تحت روش رگرسیون وزن‌دار جغرافیایی برای ۱۸۵ ایستگاه هواشناسی در استان گیلان و مجاورت آن را مورد بررسی قرار داد. آمار میانگین ده ساله بارش و موقعیت نقاط نسبت به دریا، الگوسازی شد. برای ارزیابی دقت روش رگرسیون وزن‌دار جغرافیایی و روش‌های معمول درون‌یابی شامل چندجمله‌ای سراسری و محلی، اسپلاین، وزن عکس فاصله‌ها، کریجینگ و کوکریجینگ از فن اعتبارسنجی متقابل استفاده شد و نشان داد که روش رگرسیون وزن‌دار جغرافیایی دارای کمترین مقدار خطا بوده و نتایج آن با

مواد و روش‌ها

داده‌ها و منطقه مورد مطالعه

محدوده مورد مطالعه شامل استان‌های البرز، اصفهان، تهران، زنجان، سمنان، قزوین، قم، گیلان، گلستان، مازندران، مرکزی و همدان می‌باشد (شکل ۱). اطلاعات مورد نیاز، شامل دما، رطوبت نسبی، ساعت آفتابی، تابش کل خورشیدی و بارندگی ماهانه ۵۶ ایستگاه سینوپتیک نیمه شمالی و مرکزی کشور، از ابتدای تاسیس ایستگاه‌ها تا سال ۲۰۱۰ میلادی، جمع‌آوری گردیده و پس از کنترل کیفی داده‌ها مورد استفاده قرار گرفت. در کنار داده‌های دریافتی از سازمان هواشناسی کشور، از مدل رقومی ارتفاع نیز به‌عنوان داده ورودی جهت توسعه مدل استفاده شد.

و نبود شناختی از آن‌ها باعث می‌شود تا پژوهشگران از روش‌های درون‌یابی قطعی و زمین‌آمار استفاده کنند. رگرسیون وزن‌دار جغرافیایی یکی از روش‌هایی است که می‌تواند گرادیان‌های موضعی را به خوبی شناسایی کند و در صورت وجود گرادیان، دقت درون‌یابی را افزایش دهد. در این پژوهش ابتدا مدل‌های رگرسیونی دمای هوا با فرض این که تغییرات دما تحت تاثیر عواملی مانند ارتفاع، رطوبت نسبی، ساعت آفتابی، بارندگی و تابش ورودی است به صورت تابعی از این عوامل بسط داده می‌شود. علاوه بر این روش‌های زمین‌آمار و رگرسیون وزن‌دار جغرافیایی نیز جهت درون‌یابی متغیر دما استفاده شده و در نهایت مناسب‌ترین روش جهت تعیین توزیع مکانی دما در نیمه شمالی و مرکزی کشور انتخاب می‌شود و بر مبنای آن نقشه هم‌دمای این منطقه ترسیم می‌گردد.

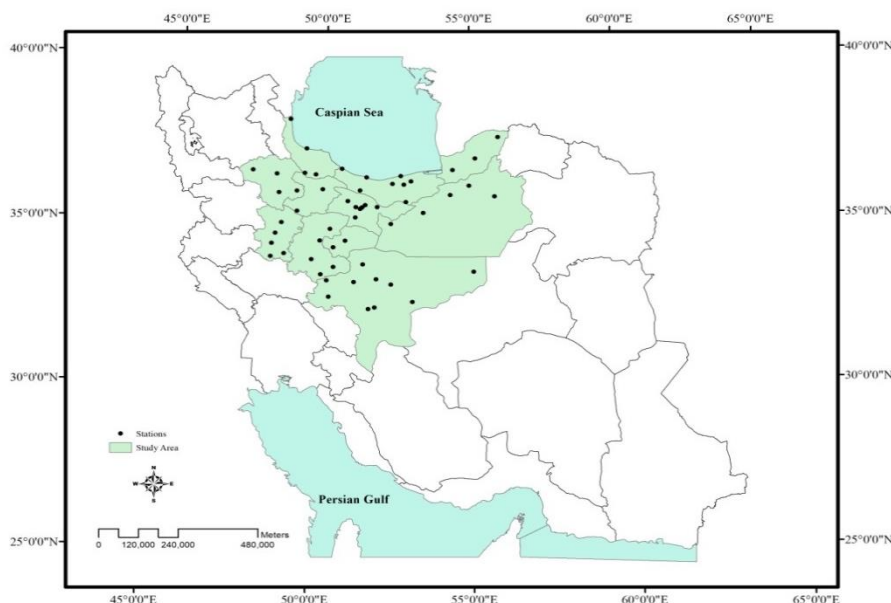


Figure 1- Location of the studied stations

شکل ۱- موقعیت ایستگاه‌های مورد بررسی

روش کریجینگ^۲

روش کریجینگ برای برآورد نقطه ناشناخته، به هر یک از نمونه‌های اندازه‌گیری شده وزنی را نسبت می‌دهد. این تخمین‌گر به افتخار یکی از پیشگامان آمار، کریجینگ نام‌گذاری شد. کریجینگ یک تخمین‌گر خطی به صورت معادله ۱ است.

$$Z^* = \sum_{i=1}^n \lambda_i \cdot Z(x_i) \quad (1)$$

که در آن Z^* مقدار متغیر مکانی برآورد شده، $Z(x_i)$ مقدار متغیر مکانی مشاهده شده در نقطه x_i و λ_i وزن

روش‌های درونیابی^۱

روش‌های مختلفی برای برآورد متغیرهایی که تغییرات زمانی و مکانی دارند، وجود دارد. تفاوت عمده این روش‌ها، محاسبه فاکتور وزنی است که به نقاط مشاهده شده اطراف نقطه مورد تخمین داده می‌شود. در این تحقیق از روش‌های کریجینگ معمولی، کوکریجینگ، رگرسیون وزن‌دار جغرافیایی و رگرسیون چند متغیره خطی به منظور درون‌یابی متغیر دما استفاده شده است.

² Kriging

¹ Interpolation

حاصل خیزی خاک و یا رطوبت خاک نیز ممکن است بر روی محصول تولیدی اثر داشته باشند. به علاوه تحلیل رگرسیون ممکن است فقط بر اساس داده‌های مشاهده شده در عرصه طبیعت نباشد، بلکه ممکن است در مطالعات آزمایشی نیز از رگرسیون چندگانه استفاده شود، که معمولاً محقق به طور همزمان تعدادی متغیر پیش‌بینی را تحت بررسی قرار می‌دهد تا بتواند اثر آن‌ها را بر روی متغیر پاسخ به دست آورد (Mesdaghi, 2004). رگرسیون وزن‌دار جغرافیایی یک روش آماری توسعه یافته جدید است، که مدل رگرسیون خطی معمولی را با برآورد پارامترهای موضعی توسعه می‌دهد. رگرسیون وزن‌دار جغرافیایی برخلاف الگوهای رگرسیون معمول که یک معادله رگرسیونی را برای توصیف روابط کلی بین متغیرها برقرار می‌کنند، اطلاعات مکانی تولید می‌کند، که تغییرات مکانی بین متغیرها را بیان می‌کند. GWR این معادلات مجزا را با استفاده از متغیر وابسته و متغیرهای توصیفی، با توجه به پهنای باند هدف، تشکیل می‌دهد. شکل و سایز پهنای باند، وابسته به ورودی‌های کاربر برای نوع کرنل^۳، نوع پهنای باند، فاصله و تعداد عارضه‌ها دارد. بنابراین نقشه‌های تولید شده از این تحلیل‌ها نقش کلیدی در توصیف غیر ایستایی مکانی بین متغیرها بازی می‌کند (Mennis, 2006). در مدل‌های رگرسیون موضعی، هنگامی که دو یا چند متغیر دارای هم خطی هستند، (دو یا چند متغیر زائد هستند و یا یک موضوع دارند)، نتایج غیر قابل اطمینان است. GWR یک معادله رگرسیون موضعی برای هر عارضه در مجموعه داده‌ها می‌سازد. هنگامی که مقادیر متغیرهای توصیفی قالب مکانی دارند، به احتمال خیلی زیاد با هم خطی موضعی به مشکل خواهند خورد. فرم کلی مدل GWR به صورت معادله ۳ است.

$$y_i = \beta_{i0} + \sum_{k=1}^m \beta_{ik} x_{ik} + \varepsilon_i \quad (3)$$

که در آن، y_i متغیر وابسته در موقعیت i ، x_{ik} ، K امین متغیر مستقل در i امین موقعیت بوده است. K امین متغیر مستقل در موقعیت i ، m تعداد متغیرهای مستقل، β_{i0} عرض از مبدا در موقعیت i ، β_{ik} ضریب سراسری رگرسیون برای k امین متغیر مستقل در موقعیت i ، و ε_i خطای تصادفی در موقعیت i است. GWR به ضرایب این اجازه را می‌دهد که به طور پیوسته در منطقه مورد مطالعه تغییر

آماري که به نمونه x_i نسبت داده می‌شود و بیان‌گر اهمیت نقطه i ام در برآورد است. این نوع کریجینگ را کریجینگ خطی گویند، زیرا ترکیبی خطی از n داده می‌باشد. شرط استفاده از این تخمین‌گر آن است که متغیر Z ، توزیع نرمال داشته باشد. در غیر این صورت باید متغیر مورد نظر را نرمال نمود و سپس از این روش استفاده کرد (Mehdizadeh, 2002). در روش کریجینگ معمولی، میانگین محلی نمونه‌ها با میانگین جامعه برابر است، ولی مقدار آن مجهول و مستقل از مختصات فرض می‌شود. بنابراین فقط از نمونه‌های محلی برای برآورد مقادیر مجهول استفاده می‌شود (Nadi, 2010). در برخی موارد ممکن است از یک متغیر به اندازه کافی نمونه‌برداری نشده باشد بنابراین بر اساس نمونه‌های موجود نمی‌توان تخمین را با دقت مورد نظر انجام داد. در چنین مواردی می‌توان با استفاده از روش کوکریجینگ و در نظر گرفتن رابطه مکانی بین متغیر مورد نظر و متغیر کمکی که داده‌های آن موجود می‌باشد، تخمین را اصلاح کرد. البته این روش زمانی به کار می‌رود که ضریب همبستگی بین دو متغیر، معنی‌دار و قوی باشد ($r \geq 0.5$). معادله کوکریجینگ به صورت معادله ۲ است.

$$Z^*(x_i) = \sum_{i=1}^n (\lambda_i \cdot Z(x_i)) + \sum_{k=1}^n (\lambda_k \cdot Z(x_k)) \quad (2)$$

که در آن Z^* مقدار تخمین زده شده برای نقطه x_i ، λ_i وزن مربوط به متغیر Z ، λ_k وزن مربوط به متغیر کمکی، $Z(x_i)$ مقدار مشاهده شده متغیر اصلی و $Z(x_k)$ مقدار مشاهده شده متغیر کمکی است. در این روش برای تخمین از نیم‌تغییرنمای متقابل^۱ استفاده می‌شود (Madani, 1994).

رگرسیون وزن‌دار جغرافیایی^۲

تحلیل رگرسیون روشی آماری است که در آن از رابطه بین دو یا چند متغیر کمی استفاده می‌شود، تا یک متغیر، توسط متغیر یا متغیرهای دیگر پیش‌بینی گردد. مدل رگرسیون ساده فقط با یک متغیر پیش‌بینی x مورد بررسی قرار می‌گیرد. در بسیاری از مطالعات میدانی علوم زیستی و کشاورزی، معمولاً بیش از یک متغیر پیش‌بینی روی متغیر پاسخ تاثیر می‌گذارد. مثلاً فقط میزان بارندگی تعیین کننده تولید نیست و عوامل دیگری مانند

¹ Cross variogram

² GWR: Geographically Weighted Regression

³ Kernel

که در آن‌ها $\tilde{Z}(x_i)$ مقدار برآورد شده متغیر x_i ، $Z(x_i)$ مقدار مشاهده شده متغیر x_i و n تعداد داده‌ها است. مناسب‌ترین روش، روشی است که کمترین MAE، MBE و RMSE را داشته باشد. هرچه MAE به صفر نزدیک‌تر باشد، بیانگر این است که مقادیر برآورد شده توسط مدل، به مقادیر حقیقی نزدیک‌تر و در نتیجه مدل از خطای کمتری برخوردار است و هرچه MBE به صفر نزدیک‌تر باشد، بیانگر این مطلب است که متوسط مقادیر مشاهده شده و برآورد شده اختلاف کمتری دارند و در نتیجه مدل، فضای مورد مطالعه را با خطای کمتری برآورد می‌کند. همچنین به منظور بیان معنی‌دار بودن اختلاف میانگین‌ها در روش‌های استفاده شده، از آزمون t-student استفاده شده است.

تجزیه و تحلیل آماری و بررسی نرمال بودن داده‌ها

پیروی داده‌ها از توزیع نرمال یکی از شرایط اصلی استفاده از روش کریجینگ است، به این منظور با انجام آزمون آماری کلموگروف-اسمیرنوف و محاسبه P-Value حاصل از این آزمون، نرمال بودن داده‌ها مورد بررسی قرار گرفت. در این رابطه داده‌هایی که از توزیع نرمال پیروی نمی‌کردند، یا به عبارت دیگر، P-Value به دست آمده از آن‌ها در آزمون مورد نظر، کمتر از ۰/۰۵ بود، با تبدیلی مناسب به توزیع نرمال تبدیل شدند. به منظور تجزیه و تحلیل آماری داده‌های دما، مقادیر میانگین، انحراف معیار، ضریب تغییرات، ضریب کشیدگی و ضریب چولگی متوسط ماهانه و سالانه داده‌های ۵۶ ایستگاه دماسنجی در منطقه مورد مطالعه محاسبه شد. نتایج حاصل از این بررسی در جدول ۱ ارائه شده است. بر اساس مقادیر ذکر شده در جدول ماه ژوئیه با میانگین دمای ۲۷/۳۲ درجه سانتی‌گراد و انحراف معیار ۳/۳۱ درجه سانتی‌گراد گرم‌ترین و ژانویه با متوسط ۲/۷۶ درجه سانتی‌گراد و انحراف معیار ۳/۵۳ درجه سانتی‌گراد سردترین ماه سال در منطقه مطالعاتی است. ژانویه با ضریب تغییرات ۱۲۷ درصد بیش‌ترین ضریب تغییرات را دارد. بررسی ارتفاع ایستگاه‌های دماسنجی منطقه به عنوان یکی از مهم‌ترین عوامل موثر بر دما، نشان می‌دهد که متوسط ارتفاع ایستگاه‌ها ۱۲۲۵ متر و حداقل ارتفاع ایستگاه‌ها ۲۳- متر و حداکثر آن ۲۴۶۵ متر می‌باشد. جدول ۲ نتایج حاصل از این بررسی را نشان می‌دهد.

کنند، و مجموعه‌ای از ضرایب در هر موقعیت می‌تواند تخمین زده شود. مشاهدات نزدیک اثر بیشتری در تخمین مجموعه ضرایب نسبت به مشاهدات دورتر دارند (Fotheringham, 2003). در اصل، GWR روابط اساسی اطراف هر نقطه رگرسیون i را اندازه‌گیری می‌کند، در حالی که هر مجموعه‌ای از ضرایب رگرسیون توسط حداقل مربعات برآورد می‌شود. ضرایب علی‌رغم فرض ثابت بودن، نسبت به مکان i حساس هستند. کالیبراسیون مدل GWR، به توالی رگرسیون‌های خطی موضعی برای برآورد هر نقطه در فضا با استفاده از زیر نمونه‌های وزن‌دار داده‌ها از مشاهدات نزدیک، بستگی دارد. مدل برای هر مشاهده i ، ماتریس وزن‌ها را که در آن مقادیر بزرگتر به نزدیکترین مشاهده i تعلق می‌گیرد، محاسبه می‌کند. رگرسیون وزن‌دار جغرافیایی برای برآورد ضرایب خطی موضعی استفاده می‌شود و متغیرهای وابسته را برآورد می‌کند.

معیارهای ارزیابی

برای ارزیابی دقت روش‌های به کار رفته، از تکنیک اعتبارسنجی متقابل استفاده شده است. در این تکنیک، یک نقطه به طور موقت حذف می‌شود و با استفاده از سایر نقاط، و روش میان‌یابی مورد نظر، مقدار نقطه مذکور محاسبه می‌شود. سپس مقدار حقیقی آن نقطه به محل خود برگردانده شده و این عمل برای نقطه بعدی تکرار می‌شود و به این ترتیب برای کلیه نقاط برآورد صورت می‌گیرد. در نهایت می‌توان نتیجه حاصل از این محاسبات را به صورت جدولی که دارای ستونی از مقادیر مشاهده شده و ستونی دیگر از مقادیر برآورد شده است، ارائه نمود، و با مقایسه این دو ستون بر اساس معیارهای پیش‌بینی روش‌ها ارزیابی می‌شوند. شاخص‌های به کار رفته برای این منظور میانگین خطا (MAE)^۱، میانگین خطای ارببی (MBE)^۲ و ریشه مجذور مربعات خطا (RMSE)^۳ با استفاده از معادله‌های ۴ تا ۶ محاسبه می‌شود.

$$MAE = \sum_{i=1}^n |\tilde{Z}(x_i) - Z(x_i)| / n \quad (4)$$

$$MBE = \sum_{i=1}^n (\tilde{Z}(x_i) - Z(x_i)) / n \quad (5)$$

$$RMSE = \left(\sum_{i=1}^n (\tilde{Z}(x_i) - Z(x_i))^2 / n \right)^{1/2} \quad (6)$$

¹ MAE: Mean Absolute Error

² MBE: Mean Bios Error

³ RMSE: Root Mean Square Error

Table 1- The statistical parameters of mean monthly air temperature (°C)

جدول ۱- پارامترهای آماری متوسط دمای ماهانه (درجه سانتی گراد)

Month	Average	Minimum	Maximum	Variation coefficient	Skewness coefficient	Kurtosis coefficient
Jan.	2.76	-4.61	8.71	127.67	-0.14	-0.79
Feb.	4.92	-2.87	9.55	65.47	-0.63	-0.41
Mar.	9.32	0.98	14.75	30.53	-0.58	0.24
Apr.	14.56	7.21	21.17	20.79	-0.12	-0.28
May	19.43	12.09	26.51	17.08	-0.08	-0.50
Jun.	24.54	17.16	31.24	13.94	-0.15	-0.49
Jul.	27.32	18.60	33.27	12.11	-0.38	0.06
Aug.	26.87	19.92	31.64	10.70	-0.44	-0.09
Sep.	22.84	16.05	27.42	12.37	-0.52	-0.35
Oct.	17.24	10.35	21.41	15.80	-0.64	-0.32
Nov.	10.15	3.63	14.34	28.95	-0.33	-0.84
Dec.	5.23	-1.25	10.27	58.52	-0.05	-0.68

Table 2- The Statistical parameters of stations' altitude

جدول ۲- پارامترهای آماری ارتفاع ایستگاهها

Average	Minimum	Maximum	Variation coefficient	Skewness coefficient	Kurtosis coefficient
1225.6	-23.6	2465.2	58.09	-0.56	-0.62

بحث و نتایج

همبستگی ساده خطی دما با متغیرهای محیطی

بیشترین ضریب همبستگی مربوط به ارتفاع، در ماههای آوریل و مه بارندگی و در ماههای ژوئن و ژوئیه بیشترین ضریب همبستگی مربوط به متغیر رطوبت نسبی است. لذا دما بیشترین همبستگی را با ارتفاع دارد.

مقادیر ضریب همبستگی دما با سایر متغیرها در جدول ۳ نشان داده شده است. بر این اساس در ماههای ژانویه، فوریه، مارس، اوت، سپتامبر، اکتبر، نوامبر و دسامبر

Table 3- Correlation coefficients of temperature with environmental parameters

جدول ۳- ضرایب همبستگی دما با پارامترهای محیطی

Month	Variable						
	Elevation	Relative Humidity	Sunshine Hour	Precipitation	Solar radiation	Latitude	longitude
Jan.	-0.81	-0.19	-0.19	0.27	-0.78	0.24	0.47
Feb.	-0.69	-0.04	0.16	-0.06	-0.66	0.05	-0.48
Mar.	-0.48	-0.19	-0.13	-0.42	-0.45	-0.14	0.42
Apr.	-0.36	-0.32	-0.09	-0.74	-0.3	0.23	0.48
May	-0.37	-0.31	-0.13	-0.61	-0.32	-0.2	0.49
Jun.	-0.26	-0.39	-0.19	-0.26	-0.23	-0.29	0.42
Jul.	-0.2	-0.46	-0.2	-0.36	-0.16	-0.35	0.34
Aug.	-0.32	-0.27	-0.18	-0.16	-0.28	-0.21	0.36
Sep.	-0.47	-0.082	-0.062	-0.002	-0.45	-0.095	0.43
Oct.	-0.61	0.085	-0.126	0.13	-0.59	0.041	0.42
Nov.	-0.78	0.24	-0.18	0.33	-0.74	0.19	0.44
Dec.	-0.82	0.32	-0.19	0.38	-0.18	0.25	0.40

تحلیل روش رگرسیون چندگانه خطی دما با سایر پارامترها به روش گام به گام

جدول ۴ میزان خطای حاصل از روش رگرسیون چندگانه خطی و متغیرهای موثر را نشان می‌دهد. همان‌طور که در جدول آمده است، متغیرهای ارتفاع، رطوبت نسبی و تابش ورودی در پیش‌بینی دما در تمامی ماه‌ها استفاده شده‌اند و بیشترین اثر را در این روش داشته‌اند. در این جدول، Z، ارتفاع، RH، رطوبت نسبی، P، بارندگی، S، تابش ورودی دریافتی، Lat، عرض جغرافیایی و Lon، طول جغرافیایی است.

در روش رگرسیون چندگانه خطی، متغیرهای رطوبت نسبی، بارندگی، ساعت آفتابی، تابش ورودی خورشید، ارتفاع، طول و عرض جغرافیایی را به مدل اضافه کرده و زیرمجموعه‌های دو، تا هفت‌تایی را مشخص نموده و در نهایت بهترین زیرمجموعه را که بالاترین ضریب تبیین را داراست، به عنوان بهترین متغیرها برای پیش‌بینی دمای هوا انتخاب و معادله رگرسیون خطی آن را به‌دست آمد.

Table 4- Error criteria in multiple linear regression
جدول ۴- معیارهای خطا در روش رگرسیون چندگانه خطی

Month	Regression variables	RMSE (°C)	MBE (°C)	R ² (%)
Jan.	Z, RH, P, S, Lat, Lon	1.78	-0.11	80
Feb.	Z, RH, N, P, S, Lat	1.65	-0.14	79
Mar.	Z, RH, N, S, Lat	1.94	-0.11	73
Apr.	Z, RH, P, S	2.36	-0.32	78
May	Z, RH, P, S	2.67	-0.33	76
Jun.	Z, RH, S, Lat, Lon	2.92	-0.23	73
Jul.	Z, RH, S, Lon	2.48	-0.18	74
Aug.	Z, RH, S, Lon	2.60	-0.25	69
Sep.	Z, RH, S, Lon	2.43	-0.22	65
Oct.	Z, RH, S	1.86	-0.15	65
Nov.	Z, RH, p, S	1.68	-0.10	75
Dec.	Z, RH, P, S, Lat	2.47	0.33	79

در ماه ژانویه، ضریب تبیین در مدل رگرسیون چندگانه خطی ۸۰٪ است و متغیر تابش ورودی دریافتی با ضریب تبیین ۵۹٪ کمترین اثر را بر مدل دما داشته است. در ماه فوریه، ضریب تبیین مدل رگرسیون چندگانه خطی به ۷۹٪ می‌رسد. در این مدل بیشترین اثر را زیرمجموعه دوتایی ارتفاع و رطوبت نسبی با ضریب تبیین ۷۵٪ داشته‌اند. در ماه مارس، ضریب تبیین مدل رگرسیون چندگانه خطی ۷۳٪ برآورد شده است و تابش ورودی دریافتی، با ضریب تبیین ۲۲٪ کمترین اثر را در مدل دما داشته است. در ماه آوریل، ضریب تبیین مربوط به مدل رگرسیون چندگانه خطی برابر با ۷۸٪ و کمترین اثر مربوط به متغیر عرض جغرافیایی با ضریب تبیین ۱۹٪ می‌باشد. ضریب تبیین برآورد دما در مدل رگرسیون چندگانه خطی ماه مه ۷۶ است. کمترین اثر را در این ماه متغیرهای ارتفاع و سپس طول جغرافیایی دارند. ضریب تبیین مدل رگرسیون چندگانه خطی ماه ژوئن، ۷۳٪ برآورد شد. کمترین تاثیر را در این ماه، مجموع دوتایی متغیرهای رطوبت نسبی و طول جغرافیایی داشته‌اند. ضریب تبیین مدل رگرسیون چندگانه خطی ماه ژوئیه به ۷۴٪ می‌رسد. این مقدار در حالتی است که متغیرهای ارتفاع، تابش ورودی، طول جغرافیایی، ساعت آفتابی و رطوبت در نظر گرفته می‌شوند. کمترین تاثیر را ابتدا رطوبت نسبی و سپس بارندگی داشته است. ضریب تبیین در مدل رگرسیون چندگانه خطی در ماه اوت، برابر با ۶۹٪ است. طول جغرافیایی و ارتفاع در این مدل کمترین اثرگذاری را داشته‌اند. در ماه سپتامبر، ضریب تبیین در مدل رگرسیون چندگانه خطی ۶۵٪ است. کمترین اثر را در این ماه، متغیرهای ارتفاع و تابش در مدل دارند. در ماه دهم، یعنی ماه اکتبر نیز ضریب تبیین ۶۵٪ است. کمترین

اثر را متغیرهای تابش ورودی و ارتفاع بر مدل داشته‌اند. در ماه نوامبر، ضریب تبیین مدل رگرسیون چندگانه خطی برابر با ۷۵٪ است. در این ماه نیز کمترین اثر را متغیرهای ارتفاع و تابش ورودی بر مدل داشته‌اند. در ماه دسامبر، ضریب تبیین مدل رگرسیون چندگانه خطی برابر با ۷۹٪ است. مانند ماه قبل، کمترین اثر را باز هم متغیرهای ارتفاع و تابش ورودی با ضریب تبیین ۶۷٪ به خود اختصاص داده‌اند. طبق نتایج به دست آمده ماه‌های فوریه و نوامبر با کمترین خطا، بهترین پیش‌بینی را داشته‌اند و بیشترین ضریب تبیین مربوط به ماه ژانویه با پارامترهای ارتفاع، رطوبت نسبی، بارندگی، تابش خورشید، طول و عرض جغرافیایی برابر با ۸۰٪ است.

ارزیابی روش‌های زمین‌آمار

در ادامه، نتایج حاصل از روش‌های مختلف زمین‌آمار، برای برآورد دمای متوسط ماهانه در منطقه مورد مطالعه به طور کامل بیان شده است.

کریجینگ معمولی

در روش کریجینگ، باید داده‌ها از توزیع نرمال پیروی نمایند، لذا قبل از انجام این روش، نرمال بودن داده‌های متوسط دمای ماهانه بررسی گردید و نمودار توزیع فراوانی آن‌ها رسم شد. نتیجه این بررسی حاکی از آن است که به غیر از ماه‌های فوریه و اکتبر، سایر ماه‌ها از توزیع نرمال پیروی می‌کنند. برای نرمال کردن داده‌ها می‌توان از توابع تبدیل مختلف نظیر لگاریتم‌گیری، استفاده از جذر داده‌ها و توابع دیگر استفاده کرد، که انتخاب تابع مناسب با سعی و خطا و بررسی نرمال بودن داده‌های تبدیل شده میسر است. پس از بررسی و تبدیل داده‌ها، با استفاده از نرم‌افزار GS+ نیم‌تغییرنمای تجربی و مدل مناسب برای دما در ماه‌های مختلف رسم گردید. در این نرم‌افزار شش نوع مدل نیم‌تغییرنمای شامل مدل‌های کروی^۱، نمایی^۲، خطی^۳، خطی دارای حد آستانه^۴ و گوسی^۵ بر داده‌ها برازش داده شده و از بین آن‌ها نیم‌تغییرنمای بهینه بر اساس مقدار مجموع مربعات خطا (RSS)^۶ کمتر و ضریب تبیین (R²) بالاتر انتخاب شد. سپس، دقت و انحراف مدل با استفاده از

¹ Spherical

² Exponential

³ Linear

⁴ Linear to sill

⁵ Gaussian

⁶ Residual Sums of Squares

می‌کنند. همچنین نسبت اثر قطعه‌ای به آستانه دارای میانگین ۴۰ درصد می‌باشد که از ۱۹ درصد در آوریل تا ۵۰ درصد در ماه‌های اکتبر و مارس متغیر است و این نشان‌دهنده ساختار مکانی قوی در داده‌ها است. همچنین میانگین دامنه تاثیر ۱۴/۵۵ کیلومتر است، که نسبت به روش کریجینگ تغییرات زیادی نداشته است و از ۷ کیلومتر در دسامبر تا ۲۱ کیلومتر در مارس متغیر است.

روش رگرسیون وزن دار جغرافیایی

رگرسیون وزن دار جغرافیایی بر خلاف الگوهای رگرسیون معمول که یک معادله رگرسیونی را برای توصیف روابط کلی بین متغیرها برقرار می‌کنند، اطلاعات مکانی تولید کرده، که تغییرات مکانی بین متغیرها را بیان می‌کند. برای استفاده از این روش، با استفاده از نرم‌افزار ArcGIS، رابطه بین متغیرهای مختلف بررسی شده و با استفاده از ضریب تبیین آن‌ها، بررسی می‌شود که کدام متغیرها برای درون‌یابی دما در هر ماه مناسب هستند. در مدل‌های رگرسیون موضعی، هنگامی که دو یا چند متغیر دارای هم‌خطی هستند، (دو یا چند متغیر زائد هستند و یا یک موضوع دارند)، نتایج غیر قابل اطمینان است. پهنای باند، پارامتر کنترل‌کننده کلیدی است که می‌تواند توسط فاصله ثابت (پهنای باند ثابت) و یا تعداد ثابت همسایگی‌های نزدیک (پهنای باند تطبیقی) مشخص شود. به منظور بررسی متغیرهای مختلف، باید نوع هسته و پهنای باندهای مختلف را مورد آزمون قرار داده و از بین آن‌ها مدلی را که بیشترین ضریب تبیین را دارد، انتخاب شوند. جدول ۸ نوع هسته، روش تعیین پهنای باند، متغیرهای مورد استفاده و ضریب تبیین در ماه‌های مختلف را نشان می‌دهد. لازم به ذکر است، در ماه‌هایی که دو متغیر رطوبت نسبی و ارتفاع بیشترین تاثیر را داشته‌اند، ضریب تبیین اختلاف ناچیزی با حالت انتخاب هر سه متغیر بالا داشته است. با این توضیح، می‌توان اظهار کرد که در تمامی ماه‌ها، هر سه متغیر ارتفاع، رطوبت نسبی و ساعت آفتابی بیشترین اثر را در مدلسازی دما داشته‌اند و بهترین نوع هسته، تطبیقی، با پهنای باند تعیین شده توسط کاربر است. با توجه به RMSE به دست آمده در ماه‌های مختلف، می‌توان اظهار کرد، ماه دسامبر با کمترین خطا بهترین نتیجه را در پیش‌بینی دما داشته است. بیشترین خطا نیز در ماه ژوئن مشاهده شده است.

عمل اعتبارسنجی حذفی محاسبه شد. جدول ۵ نتیجه ارزیابی روش کریجینگ معمولی و پارامترهای نیم‌تغییرنمای تجربی برای داده‌های دما را در منطقه مطالعاتی نشان می‌دهد.

کوکریجینگ

زمانی که از داده‌های متغیر اصلی به تعداد کافی موجود نباشد و یک متغیر کمکی در منطقه باشد که به تعداد بیشتری در منطقه موجود باشد و با متغیر اصلی همبستگی قوی داشته باشد، از این روش استفاده می‌شود اما شرط استفاده از این روش آن است که، قدر مطلق ضریب همبستگی بین متغیر اصلی و متغیر کمکی، معنی‌دار و بزرگتر از ۰/۵ باشد. به همین منظور، رابطه بین دما در ماه‌های مختلف سال با متغیرهای دیگر، همچون ارتفاع، رطوبت نسبی، بارندگی و غیره بررسی گردید. نتیجه این بررسی نشان داد که در هشت ماه از سال، متوسط دما با ارتفاع ایستگاه‌ها، در دو ماه از سال با بارندگی و در دو ماه دیگر با رطوبت نسبی دارای همبستگی بیش از ۵۰ درصد است. جدول ۶ ضرایب همبستگی دما با این متغیرها به علاوه معادلات همبستگی آن‌ها را در ماه‌های مختلف نشان می‌دهد. همان‌طور که در جدول ۶ دیده می‌شود، بیشترین همبستگی را متغیر کمکی ارتفاع با دما دارد و دمای دسامبر بیشترین همبستگی را با ارتفاع دارد، همچنین دمای اوت دارای کمترین همبستگی با ارتفاع می‌باشد. در ماه‌های آوریل و مه، ضریب همبستگی نشان از همبستگی بالای این متغیر با دما دارند، اما در ماه‌های ژوئن، ژوئیه، اوت و سپتامبر، قدر مطلق ضریب همبستگی کمتر از ۰/۵ است، و به همین دلیل در این چهار ماه از روش کوکریجینگ استفاده نشده است که در جدول بعد با خط‌چین مشخص گردیده است. در این روش، تغییرات همزمان دو متغیر اصلی و کمکی (نیم‌تغییرنمای متقابل)^۱ نیز باید ساختاردار باشد. پس از تعیین متغیر کمکی و ضریب همبستگی آن با دمای هوا، نیم‌تغییرنمای متقابل متغیر کمکی با دما در ماه‌های مختلف رسم گردید و با عمل اعتبارسنجی حذفی، دقت و انحراف مدل محاسبه شد که حاصل آن در جدول ۷ خلاصه شده است. همان‌طور که در این جدول دیده می‌شود، ۷۵ درصد نیم‌تغییرنماها از مدل کروی پیروی

^۱ Cross-Variogram

Table 5- Optimal Semivariogram characteristics and the results of the Kriging method evaluation for monthly mean temperature

جدول ۵- مشخصات نیم تغییرنمای بهینه و نتایج ارزیابی روش کریجینگ برای داده‌های دمای متوسط ماهانه

Month	مدل	Nugget (C ₀)	Sill=C ₀ +C	Range	C/Sill	C ₀ /Sill	RMSE (°C)	MBE (°C)
Jan.	Exponential	4.47	16.78	3.30	0.61	0.39	2.60	-0.67
Feb.	Exponential	7.56	20.66	13.69	0.63	0.37	2.51	-0.60
Mar.	Exponential	6.98	13.97	20.99	0.50	0.50	2.44	-0.34
Apr.	Spherical	6.97	16.18	19.71	0.57	0.43	2.62	-0.44
May	Spherical	8.30	20.10	20.36	0.59	0.41	2.83	-0.47
Jnu.	Exponential	10.10	20.15	20.99	0.50	0.50	2.74	-0.34
Jul.	Linear	10.65	10.65	6.38	0.00	1.00	2.56	-0.24
Aug.	Linear	8.01	8.01	6.38	0.00	1.00	2.39	-0.37
Sep.	Linear	7.71	7.98	6.38	0.03	0.97	2.38	-0.47
Oct.	Exponential	6.39	12.78	20.99	0.50	0.50	2.38	-0.49
Nov.	Exponential	6.74	15.46	13.43	0.56	0.44	2.43	-0.60
Dec.	Exponential	6.37	19.88	13.11	0.68	0.32	2.36	-0.70

Table 6- Coefficient of determination and Relationship between monthly temperature and auxiliary variable

جدول ۶- ضریب تبیین و رابطه بین دمای ماهانه و متغیر کمکی

Month	Auxiliary variable	Correlation equation	Coefficient of determination	P-Value
Jan.	Altitude	T=7.692-0.004Z	-0.81	0.00
Feb.	Altitude	T=8.797-0.003Z	-0.70	0.00
Mar.	Altitude	T=11.67-0.002Z	-0.48	0.00
Apr.	Precipitation	T=19.42-0.113Z	-0.74	0.00
May	Precipitation	T=23.23-0.151Z	-0.61	0.00
Jnu.	Relative Humidity	T=27.13-0.067Z	-0.39	0.003
Jul.	Relative Humidity	T=30.32-0.078Z	-0.46	0.00
Aug.	Altitude	T=28.49-0.001Z	-0.32	0.015
Sep.	Altitude	T=25.12-0.002Z	-0.47	0.00
Oct.	Altitude	T=20.12-0.002Z	-0.61	0.00
Nov.	Altitude	T=14.11-0.003Z	-0.78	0.00
Dec.	Altitude	T=9.566-0.003Z	-0.82	0.00

Table 7- Semivariogram characteristics of the variation of the auxiliary temperature data and the results of the evaluation of the Co-Kriging method

جدول ۷- مشخصات نیم تغییرنمای متقابل داده‌های متغیر کمکی - دما و نتایج ارزیابی روش کوکریجینگ

Month	Model	Nugget (C ₀) Nugget variance	Sill=C ₀ +C	Range	C/Sill	C ₀ /Sill	RMSE (°C)	MBE (°C)
Jan.	Spherical	-1131.00	-2707.00	7.35	0.58	0.42	2.78	-0.50
Feb.	Spherical	-1158.00	-2481.00	13.95	0.53	0.47	2.96	0.40
Mar.	Exponential	-845.00	-1691.00	20.99	0.50	0.50	3.13	-0.08
Apr.	Spherical	-21.30	-113.60	17.37	0.81	0.19	3.00	-0.41
May	Spherical	-16.40	-55.07	17.06	0.70	0.30	3.20	-0.46
Jun.	-	-	-	-	-	-	-	-
Jul.	-	-	-	-	-	-	-	-
Aug.	-	-	-	-	-	-	-	-
Sep.	-	-	-	-	-	-	-	-
Oct.	Spherical	-958.00	-1920.00	18.90	0.50	0.50	2.71	-0.38
Nov.	Exponential	-1115.00	-3688.00	13.75	0.70	0.30	2.69	-0.44
Dec.	Spherical	-1035.00	-2304.00	7.02	0.55	0.45	2.38	-0.46

روش کوکریجینگ در برخی ماه‌ها به دلیل پایین بودن ضریب همبستگی متغیر کمکی، از این روش استفاده نشد. بر اساس جدول ۹، در تمامی ماه‌ها، رگرسیون وزن‌دار جغرافیایی کمترین میزان خطا را دارد.

انتخاب مناسب‌ترین روش در پهنه‌بندی دمای ماهانه برای انتخاب روش مناسب میان‌یابی دمای ماهانه، نتایج میانگین مربعات خطای روش‌های به کار رفته در این تحقیق، با استفاده از جدول ۹ با یکدیگر مقایسه شدند. در

Table 8- Monthly results of geographically weighted regression method

جدول ۸- نتایج ماهانه حاصل از روش رگرسیون وزن دار جغرافیایی

Month	Kernel type		Band Width Method		Effective variables			Error indices			
	Fixed	Adaptive	AICc	CV	Bandwidth parameter	Altitude	Relative humidity	Sunshine Hour	RMSE	MBE	R ²
Jan.		*		*		*	*		1.15	0.003	0.86
Feb.		*	*			*	*		1.30	-0.07	0.83
Mar.		*			*	*	*		1.32	-0.15	0.82
Apr.		*			*	*	*	*	1.06	-0.09	0.87
May		*			*	*	*	*	1.49	-0.04	0.83
Jun.		*			*	*	*	*	1.54	-0.03	0.83
Jul.		*			*	*	*	*	1.39	-0.08	0.84
Aug.		*			*	*	*	*	1.38	-0.004	0.79
Sep.		*			*	*	*	*	1.38	-0.01	0.78
Oct.		*			*	*	*	*	1.30	-0.06	0.79
Nov.		*			*	*	*	*	0.95	-0.08	0.86
Dec.		*			*	*	*	*	0.83	-0.02	0.87

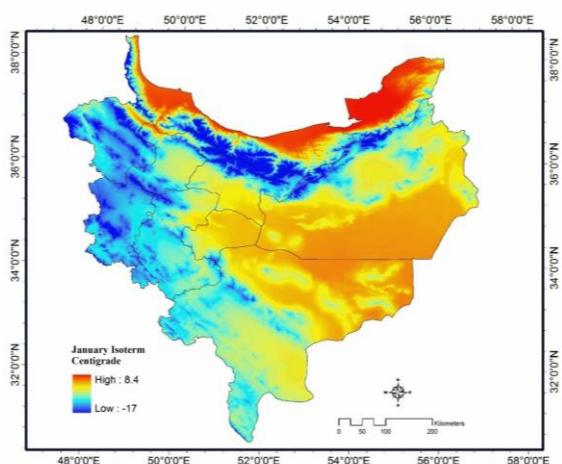


Figure 2- January temperature zoning map using geographical weighting regression

شکل ۲- نقشه پهنه بندی دمای ماه ژانویه با استفاده از روش رگرسیون وزن دار جغرافیایی

نتیجه گیری

دما یکی از عناصر اساسی شناخت هوا می باشد، با توجه به دریافت نامنظم انرژی خورشیدی توسط زمین، دمای هوا در سطح زمین دارای تغییرات زیادی است که این تغییرات به نوبه خود سبب تغییرات دیگری در سایر عناصر هوا می گردد. در این تحقیق، کارایی روش های درون یابی زمین آماری مانند کریجینگ و کوکریجینگ، روش رگرسیون وزن دار جغرافیایی و رگرسیون چندمتغیره خطی بررسی و تحلیل گردید. ارزیابی بر اساس تکنیک اعتبارسنجی متقابل انجام شد. بر اساس نتایج اصلی به دست آمده در ماه های ژانویه، فوریه، مارس، اوت، سپتامبر، اکتبر، نوامبر و دسامبر بیشترین ضریب همبستگی مربوط به ارتفاع، در ماه های آوریل و مه، بازندگی و در ماه های ژوئن و ژوئیه بیشترین ضریب همبستگی مربوط به پارامتر رطوبت نسبی است. پس می توان نتیجه گرفت، دما

بنابراین این روش به عنوان بهترین روش برآورد داده های دمای ماهانه در نیمه شمالی و مرکزی ایران معرفی می گردد. پس از آن به ترتیب، روش رگرسیون چند متغیره خطی، کریجینگ و کوکریجینگ در اولویت های بعدی قرار دارند.

نقشه پهنه بندی دما در نیمه شمالی کشور

در آخر با استفاده از بهترین روش معرفی شده، یعنی روش رگرسیون وزن دار جغرافیایی، نقشه پهنه بندی دمای ماهانه را رسم شده است. در شکل ۲ نقشه پهنه بندی دمای ماه ژانویه با استفاده از روش رگرسیون وزن دار جغرافیایی ملاحظه می شود. بررسی این شکل نشان می دهد که در روش رگرسیون وزن دار جغرافیایی اثر ارتفاع بر تغییرات دما به خوبی منعکس شده است به صورتی که دماهای پایین منطبق بر مناطق مرتفع و دماهای بالا منطبق بر مناطق کم ارتفاع به ویژه در بخش جنوب شرقی ناحیه مورد مطالعه می باشند.

Table 9- The mean square error of the used methods
جدول ۹- میانگین مربعات خطا در روش های استفاده شده

Month	Multivariate linear regression	Geographical Weighted Regression	Cokriging	Kriging
Jan.	1.78	1.15	2.78	2.6
Feb.	1.65	1.3	2.96	2.51
Mar.	1.94	1.32	3.13	2.44
Apr.	2.36	1.06	3	2.62
May	2.67	1.49	3.2	2.83
Jun.	2.92	1.54	-	2.74
Jul.	2.48	1.39	-	2.56
Aug.	2.6	1.38	-	2.39
Sep.	2.43	1.38	-	2.38
Oct.	1.86	1.3	2.71	2.38
Nov.	1.68	0.95	2.69	2.43
Dec.	2.47	0.83	2.38	2.36

ماه‌های مختلف، می‌توان اظهار کرد، ماه دسامبر با کمترین خطا بهترین نتیجه را در پیش‌بینی دما داشته است. بیشترین خطا نیز در ماه ژوئن مشاهده شده است. در پایان با توجه به نتایج به دست آمده در تمامی ماه‌ها، رگرسیون وزن‌دار جغرافیایی کمترین میزان خطا را نشان می‌دهد. بنابراین این روش به عنوان بهترین روش برآورد داده‌های دمای ماهانه در نیمه شمالی و مرکزی ایران معرفی می‌گردد. نتایج این تحقیق با سایر مطالعات صورت گرفته در این زمینه همخوانی دارد. به عنوان مثال Ghorbani et al., (2012) روش‌های مختلف درون‌یابی قطعی و زمین‌آمار را با روش رگرسیون وزن‌دار جغرافیایی برای پهنه‌بندی و ترسیم نقشه‌همبارش در استان گیلان مورد ارزیابی قرار دادند و نتیجه گرفتند روش رگرسیون وزن‌دار جغرافیایی در مقایسه با روش کریجینگ به ترتیب با ریشه میانگین مربعات خطای برابر با ۱۴۷ و ۱۸۷ میلی‌متر با مقدار خطای کمتر می‌تواند در درون‌یابی و پهنه‌بندی میانگین بارش سالانه مورد استفاده قرار گیرد. Benavides et al., (2007) مدل‌های برآورد دمای هوا، شامل پنج مدل زمین‌آمار و دو مدل رگرسیونی را در منطقه آستوریاس در شمال اسپانیا مقایسه کرده و به این نتیجه رسیدند که روش‌های رگرسیونی محلی از جمله روش رگرسیون وزن‌دار جغرافیایی جزئیاتی که در روش‌های سراسری نادیده گرفته می‌شوند، تا حدی به تصویر می‌کشند. همچنین Ghorbani and Aghashariatmadari (2014) با در نظر گرفتن میانگین ۳۰ ساله دمای هوا و رطوبت نسبی ۲۴۰ ایستگاه همدیدی و اقلیم‌شناسی در ایران، نشان دادند که در مورد دمای هوا که تحت تاثیر ارتفاع می‌باشد درون‌یابی به کمک روش رگرسیون وزن‌دار جغرافیایی با دیگر روش‌های درون‌یابی اختلاف معنی‌داری دارد و بهتر می‌باشد، در حالی که این اختلاف در مورد رطوبت نسبی مشاهده نشد و این می‌تواند به دلیل نبود رابطه مشخصی بین ارتفاع و رطوبت نسبی باشد. با توجه به نتایج و روش‌های مورد استفاده در این تحقیق و مطالعات مشابه پیشنهاد می‌شود که روش‌های به کار رفته در این تحقیق، در سایر مناطق کشور نیز انجام و مورد مقایسه قرار گیرند و در صورتی که با تقسیم‌بندی منطقه از نظر ارتفاعی، رابطه مناسبی بین متغیرهای مورد مطالعه و ارتفاع حاصل شود، این روش‌ها در هر طبقه ارتفاعی انجام شود.

بیشترین همبستگی را با ارتفاع دارد. طبق نتایج به دست آمده از روش رگرسیون چند متغیره خطی با استفاده از مدل گام به گام، ماه‌های فوریه و نوامبر با کمترین خطا، بهترین پیش‌بینی را داشته‌اند و بیشترین ضریب تبیین مربوط به ماه ژانویه با پارامترهای ارتفاع، رطوبت نسبی، بارندگی، تابش خورشید، طول و عرض جغرافیایی برابر با ۸۰٪ است. در روش کریجینگ، حدود ۶۰ درصد موارد مدل تئوری نیم تغییرنما، مدل نمایی بوده است و در بقیه موارد کروی و خطی است. همچنین نسبت اثر قطعه‌ای به آستانه، دارای میانگین ۵۷ درصد می‌باشد، که از ۳۲٪ در دسامبر تا ۱۰۰٪ در ژوئیه و اوت متغیر است. ساختار مکانی دما در سه ماه ژوئیه، اوت و سپتامبر با توجه به اینکه نسبت اثر قطعه‌ای به آستانه بیشتر از ۰/۵ است، نامناسب است و کاربرد زمین‌آمار در این ماه‌ها مفید واقع نمی‌شود. همچنین با توجه به نتایج، بهترین نیم تغییرنما مربوط به ماه فوریه است، زیرا میزان اثر قطعه‌ای به آستانه، حداقل است. نتایج تعیین متغیر کمکی در روش کوکریجینگ نشان داد، در هشت ماه از سال، متوسط دما با ارتفاع ایستگاه‌ها، در دو ماه از سال با بارندگی و در دو ماه دیگر با رطوبت نسبی دارای همبستگی بیش از ۵۰ درصد است و بدین ترتیب این سه متغیر به عنوان متغیر کمکی مورد استفاده قرار گرفتند. با توجه به اینکه در روش کوکریجینگ، تغییرات همزمان دو متغیر اصلی و کمکی (نیم تغییرنمای متقابل) نیز باید ساختاردار باشد. پس از تعیین متغیر کمکی و ضریب همبستگی آن با دمای هوا، نیم تغییرنمای متقابل متغیر کمکی با دما در ماه‌های مختلف رسم گردید و نشان داد که بیشترین همبستگی را متغیر کمکی ارتفاع با دما دارد و دمای دسامبر بیشترین همبستگی را با ارتفاع دارد، همچنین دمای اوت دارای کمترین همبستگی با ارتفاع می‌باشد. در ماه‌های آوریل و مه، ضریب همبستگی نشان از همبستگی بالای این متغیر با دما دارند، اما در ماه‌های ژوئن، ژوئیه، اوت و سپتامبر، قدر مطلق ضریب همبستگی کمتر از ۰/۵ است، و به همین دلیل در این چهار ماه از روش کوکریجینگ استفاده نشده است. در روش رگرسیون وزن‌دار جغرافیایی، در تمامی ماه‌ها، هر سه متغیر ارتفاع، رطوبت نسبی و ساعت آفتابی بیشترین اثر را در مدلسازی دما داشته‌اند و بهترین نوع هسته، تطبیقی، با پهنای باند تعیین شده توسط کاربر است. همچنین با توجه به RMSE به دست آمده در

منابع

- Abtew, W., Obeysekera, J., Shin, G. 1993. Spatial Analysis for monthly rainfall in south Florida. *Water Resources Bulletin*, 29(2): 179-188.
- Benavides, R., Montes, F., Rubio, A., Osoro, K. 2006. Geostatistical modelling of air temperature in a mountainous region of Northern Spain. *Agricultural and forest meteorology*, 146: 173-188.
- Bostan, P. A., Akyürek, Z. 2007. Exploring the mean annual precipitation and temperature values over Turkey by using environmental variables. In ISPRS Joint Workshop "Visualization and Exploration of Geospatial Data". University of Applied Sciences, Stuttgart.
- Fotheringham, A. S., Brunson, C., Charlton, M. 2003. Geographically weighted regression: the analysis of spatially varying relationships. John Wiley and Sons, 268 pages.
- Georganos, S., Abdi, A. M., Tenenbaum, D. E., Kalogirou, S. 2017. Examining the NDVI-rainfall relationship in the semi-arid Sahel using geographically weighted regression. *Journal of Arid Environments*, 146: 64-74.
- Ghorbani, Kh. 2013. Geographically Weighted Regression: A Method for Mapping Isohyets in Gilan Province. *Journal of water and soil*, 26(3): 743-752. (In Farsi)
- Ghorbani, Kh., Aghashariatmadari, Z. 2014. The Effect of Local Gradients on Increasing of Climatic Data Interpolation Accuracy by Geographically Weighted Regression (Case Study: Air Temperature and Relative Humidity). *Journal of Watershed Management Research*, 5(1): 132-143. (In Farsi)
- Gundogdu, I. B., Esen, O. 2010. The importance of secondary variables for mapping of meteorological data. In International Conference on Cartography and GIS. 12-15 June. Nessebar, Bulgaria. 15-20.
- Hevesi, J. A., Istok, J. D., Flint, A. L. 1992. Precipitation estimation in mountainous terrain using multivariate geostatistics. Part I: structural analysis. *Journal of applied meteorology*, 31(7): 661-676.
- Johnson, B. A., Scheyvens, H., Khalily, M. B., Onishi, A. 2019. Investigating the relationships between climate hazards and spatial accessibility to microfinance using geographically-weighted regression. *International Journal of Disaster Risk Reduction*, 33: 122-130.
- Madani, H. 1994. Geostatistical basics. Amirkabir University publication. (In Farsi).
- Mehdizadeh, H. 2002. Evaluation of Geostatistical methods to estimate temperature and rainfall in Ourmieh lake basin. Master's thesis of Agricultural Meteorology. University of Tehran. (In Farsi)
- Mennis, J. 2006. Mapping the results of geographically weighted regression. *The Cartographic Journal*, 43(2): 171-179.
- Mesdaghi, M. 2004. Regression methods for research in agriculture and natural resources. Imam Reza University, Mashhad. 290. (In Farsi)
- Nadi, M. 2010. Using the Various Interpolation Techniques of Climatic Data for Determining the Most Important Factors Affecting the Trees Growth in the Elevated Areas of Chaharbagh, Gorgan. Master's thesis of Agricultural Meteorology. University of Tehran. (In Farsi)
- Nalder, I. A., Wein, R. W. 1998. Spatial interpolation of climatic normals: test of a new method in the Canadian boreal forest. *Agricultural and forest meteorology*, 92(4): 211-225.
- Rahimi Bandarabadi, S. 2000. Investigating the Application of Geostatistic Methods for Estimating Precipitation in Arid and Semi-Arid Areas of Southeast of Iran. Master's thesis of desertification. University of Tehran. (In Farsi)
- Tobler, W. R. 1970. A computer movie simulating urban growth in the Detroit region. *Economic geography*, 46(sup1): 234-240.
- Zhai, L., Li, S., Zou, B., Sang, H., Fang, X., Xu, S. 2018. An improved geographically weighted regression model for PM2.5 concentration estimation in large areas. *Atmospheric Environment*, 181: 145-154.
- Zhou, Q., Wang, C., Fang, S. 2018. Application of geographically weighted regression (GWR) in the analysis of the cause of haze pollution in China. *Atmospheric Pollution Research* (In Press). <https://doi.org/10.1016/j.apr.2018.12.012>

Comparative study of geo-statistical and multivariate models for air temperature interpolation in central and northern regions of Iran

Z. Aghashariatmadari¹, M. Ettehad², Kh. Ghorbani³

Received: 10/03/2018

Accepted: 10/02/2019

Abstract

Air temperature is one of the major variables required for agroclimatic classifications. For spatial zoning of temperature point observations, the interpolation approaches in which the horizontal and vertical gradients are included may be applied. In this research, the skill of Kriging, Co-Kriging, geographically weighted regression and Linear Multivariate Regression was evaluated for the interpolation of the monthly mean temperature values using the data of 56 synoptic stations located in the northern and central regions of Iran. The results of the statistical analysis indicated that the geographically weighted regression have the greatest difference with the other methods in month of December, with root mean square error (RMSE) equal to 0.83 °C, Based on the RMSE values of all months, the geographically regression method (with RMSE of 1.26°C) is the most suitable approach for temperature spatial zoning in this region. and then linear multiple regression method with RMSE of 2.24 °C, Kriging with RMSE of 2.52 °C and Cokriging with highest RMSE of 2.86 °C were ranked second to fourth, respectively. Besides, it is concluded that for high altitude areas where almost no weather station exist, the geographically weighted regression method provided the most accurate interpolated data of the air temperature.

Keywords: Geographically weighted regression, Local interpolation, Multivariable regression, Temperature



¹ Assistant professor, Irrigation and Reclamation Engineering Department, University College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran

(*Corresponding Author Email Address: zagha@ut.ac.ir)

DOI: 10.22125/agmj.2019.148887.1034

² Graduated student of agrometeorology, Irrigation and Reclamation Engineering Department, University College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran

³ Associate Professor of Agrometeorology, Department of Water Engineering, Gorgan University of Agricultural Sciences and Natural Resources