نشریه تحلیل فضایی مخاطرات محیطی، سال یازدهم، شماره ۲، تابستان ۱۴۰۳

صفحات ۱۱۷ – ۱۳۶

ارزیابی آماره‌های اندازه گیری دقت مدل و استفاده صحیح از آنها (مطالعه موردی: درون‌یابی بارش استان فارس)

علی اکبر میر شفیعی؛ دانش آموخته کارشناسی ارشد گروه جغرافیای طبیعی، دانشگاه تهران، تهران، ایران.

سعید بازگیر[1]؛ دانشیار گروه جغرافیای طبیعی، دانشگاه تهران، تهران، ایران.

قاسم عزیزی؛ استاد گروه جغرافیای طبیعی، دانشگاه تهران، تهران، ایران.

دریافت مقاله : ۱۴۰۲/۰۸/۲۷ پذیرش نهایی:۱۴۰۳/۰۶/۰۲۴

چکیده

در بسیاری از تحقیقات علمی، استفاده از آماره‌های سنجش خطا در انتخاب یک مدل و یا روش در تحلیل فضایی مخاطرات محیطی، بدون در نظرگرفتن ملاحظاتی انجام می‌شود. در این پژوهش، سنجش دقت روش‌های درون‌یابی بارش در استان فارس با هدف ارزیابی عملکرد آماره‌های پرکاربرد اندازه‌گیری خطا و ارائه توصیه‌هایی برای استفاده صحیح از آنها بوده است. درون‌یابی بارش به کمک ۱۶۱ ایستگاه هواشناسی (۲۲ ایستگاه همدیدی و ۱۳۹ ایستگاه باران‌سنجی) برای سال ۱۳۹۸ به عنوان یک سال پربارش با روش‌های وزنی عکس فاصله، کریجینگ، کوکریجینگ و تابع پایه شعاعی انجام شد. نتایج ارزیابی آماره MBE نشان می‌دهد در مواردی به دلیل صفر شدن حاصل جمع مقادیر مثبت و منفی، محقق در انتخاب روش درون‌یابی دچار اشتباه شده و این آماره تنها بیش‌برآوردی و یا کم‌برآوردی را نشان داده و نمی‌توان از آن برای ارزیابی دقت و انتخاب روش‌های درون‌یابی استفاده نمود. در خصوص ضریب تبیین (r^2)، نتایج نشان داد که به دلیل عدم سنخیت بزرگی دامنه تغییرات این ضریب (صفر تا ۱) با مقادیر خطا (۱۰۰ تا ۴۰۰ میلی‌متر برای درون‌یابی بارش استان فارس)، استفاده از آن در ارزیابی دقت یک روش صحیح نمی‌باشد. درخصوص NRMSE نتایج نشان داد که در نمونه‌های با تعداد افراد کم (n=3, NRMSE=0.35) در مقایسه با نمونه‌های با تعداد زیاد (n=20, NRMSE=0.097)، مقدار آن بیش از اندازه افزایش یافته و استفاده از این آماره پیشنهاد نمی‌شود. در یک نتیجه‌گیری کلی، استفاده از آماره‌های MAE و RMSE برای ارزیابی خطای روش‌های درون‌یابی به دلیل واقعی‌تر نشان دادن مقدار خطا پیشنهاد می‌شوند.

واژه‌های کلیدی: بارش، آماره‌های ارزیابی خطا، روش‌های درون‌یابی، استان فارس

[1] Email: sbazgeer@ut.ac.ir نویسنده مسئول

[Downloaded from system.khu.ac.ir on 2026-11-03]

[DOR: 20.1001.1.24237892.1403.11.2.7.9]

[DOI: 10.61186/jsaeh.11.2.117]

مقدمه

تخمین دقیق خصوصیات کمی و کیفی پدیده‌های طبیعی لازمه صرف زمان و خرج هزینه‌های فراوانی می‌باشد. در این خصوص تشکیل مدل‌های آماری روندی کارآزموده و کارآمد می‌باشد که با تخصیص داده‌های نقطه‌ای به یک صفحه‌ی پیوسته، صرفه جویی در زمان و هزینه‌ها را مهیا کرده تا سبب شود در راستای مدیریت منابع، قدمی مهم و گران‌بها برداشته شود (زند کریمی، ۱۳۹۸). روش‌های درون‌یابی به منظور پهنه‌بندی پتانسیل خطر رخداد مخاطرات طبیعی نقش مؤثری برای برنامه‌ریزی کاهش خسارت ناشی از این مخاطرات را دارند (Kalo et al., ۲۰۲۰). روش‌های درونیابی جهت پهنه‌بندی داده‌های نقطه‌ای و تخمین مقادیر در مکان‌های فاقد داده، می‌باشد (Yuval et al., ۲۰۱۷). روش‌های جدیدی نیز برای درون‌یابی در حوزه‌های مختلفی توسعه داده شده‌اند. از جمله‌ی این روش‌ها شبکه عصبی و محاسبات فازی می‌باشد (Alimissis et. al., 2018). خوانندگان محترم می‌توانند برای جزئیات بیشتر در خصوص روش‌های درون‌یابی به مقاله ندیری و همکاران، ۱۳۹۳، مراجعه نمایند.

انتخاب دقیق‌ترین روش درون‌یابی مشروط بر داشتن کمترین خطا می‌باشد (حسینی و همکاران، ۱۴۰۲). با بکارگیری صحیح پهنه‌بندی مخاطرات محیطی مانند سیلاب می‌توان به توسعه راه حل راهبردی مناسب به منظور کاهش خطر سیل کمک کرد (خسروی و همکاران، ۱۴۰۱). برای نمونه پایش خشکسالی در هر منطقه یکی از ارکان مهم در مدیریت منابع آب می‌باشد. در این راستا درونی‌یابی یکی از روش‌های مناسب برای پایش مکانی خشکسالی است. برای ارزیابی و مقایسه مقادیر تخمینی با مشاهدات مستقل می‌توان از آماره‌های اندازه گیری خطا به صورت وسیع‌تری استفاده کرد، (Willmott, ۲۰۰۵). از آماره‌های ارزیابی خطا جهت مشخص کردن اینکه کدام روش بهترین برآورد را ارائه می‌دهد، استفاده شده است (۲۰۱۰ ,.Falivene et.al). برای بدست آوردن مقدار تخمینی تمام نقاط موجود از روش اعتبارسنجی متقابل (cross-validation)، برپایه حذف موقت یک نقطه و تخمین توسط دیگر نقاط موجود، استفاده می‌شود. بر این اساس روش کنارگذاری یک نقطه و برآورد توسط نقاط دیگر بر روی تمامی نقاط به نوبت اعمال می‌شود تا تخمین یا به عبارتی برآورد برای تمامی نقاط صورت گیرد (Davies et.al., ۱۹۸۴).

اگر نقاط به روش دسته‌ای حذف گردند اعتبارسنجی[۲] و در صورتی که نقاط یک به یک حذف گردند اعتبارسنجی متقابل[۳] نامیده می‌شود (Esri, 2020). انتخاب الگوی مناسب درون‌یابی از نقش بسزایی برخوردار می‌باشد. در این راستا طی مطالعه‌ای (Nekoamal and Mirabbasi, ۲۰۱۷) دقت دو روش وزن‌دهی عکس فاصله و کریجینگ معمولی به منظور درون‌یابی داده های سطح آب‌های زیرزمینی دشت سرخون در استان هرمزگان مورد بررسی قرار گرفت. در این تحقیق ارزیابی روش‌های درون‌یابی به کمک آماره‌های ارزیابی خطا نشان داد دقت روش کریجینگ نسبت به روش وزن‌دهی عکس فاصله، بالاتر می‌باشد. در مطالعه‌ای دیگر نشان داده شد که نتایج حاصل از روش‌های درون‌یابی در یک پدیده می‌تواند از مکانی به مکان دیگر دارای مقادیر متفاوتی باشند و انتخاب یک روش درون‌یابی به عنوان روشی برتر، تنها می‌تواند مختص همان محدوده درون‌یابی شده باشد (Garcia-Santos, ۲۰۲۰). جهت ارزیابی آماره‌های سنجش خطا پژوهش‌هایی انجام شده، که به عنوان مثال در مطالعه‌ای (Willmott, ۲۰۰۹) به بررسی میزان معناداری ضریب همبستگی پیرسون و ارزیابی معیارهای میانگین مجذور مربعات خطا

2-Validation

3 - Cross Validation

[Downloaded from system.khu.ac.ir on 2026-11-03]
[DOR: 20.1001.1.24237892.1403.11.2.7.9]
[DOI: 10.61186/jsaeh.11.2.117]

[Downloaded from system.khu.ac.ir on 2026-11-03]
[DOR: 20.1001.1.24237892.1403.11.2.7.9]
[DOI: 10.61186/jsaeh.11.2.117]

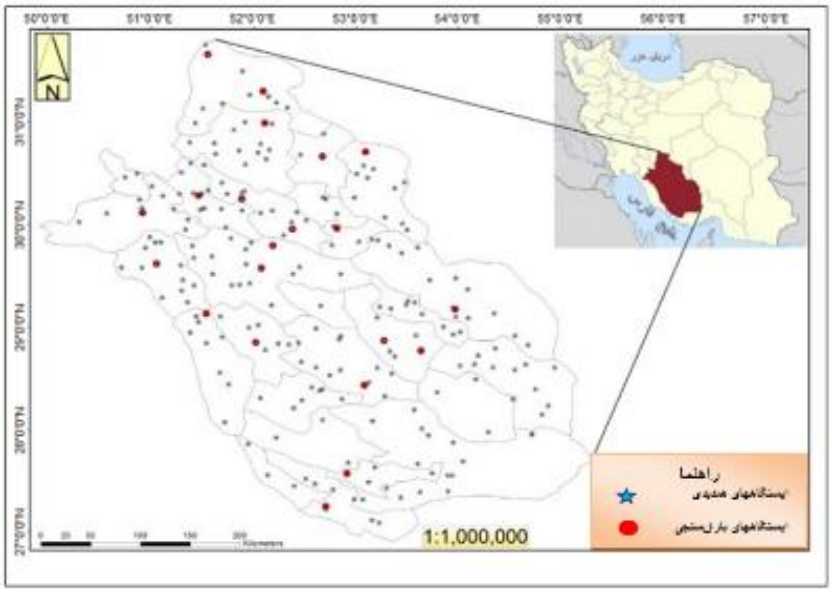
(RMSE)[۴]، میانگین خطای مطلق (MAE)[۵]، میانگین مربعات خطا (MSE)[۶]، میانگین خطای اریب (MBE)[۷] پرداخته که این اعتبارسنجی بر روی مقادیر تبخیر و تعرق انجام گرفته است. در این تحقیق نشان داده شد که ضریب همبستگی پیرسون یک آماره گمراه کننده در اندازه‌گیری دقت مدل می‌باشد. همچنین از نتایج این تحقیق می‌توان به قابل اعتماد بودن ریشه میانگین مربعات خطا، میانگین خطای مطلق و شاخص توافق (d) برای اعتبارسنجی خطا اشاره نمود. در مطالعه‌ای ارجحیت استفاده از RMSE نسبت به MAE مورد بررسی قرار گرفت (Chai et al., ۲۰۰۴). ایشان در این مطالعه نشان دادند که بر خلاف آنچه که در تحقیقات قبلی (Willmott and Matsuura, ۲۰۰۵) بیان شده است، نبود علامت قدر مطلق در RMSE از مزایای استفاده از این آماره می‌باشد. دیویس و همکاران[۸] (۱۹۸۴) در پژوهشی به این نتیجه رسیدند که انتخاب RMSE و MBE به طور جداگانه می‌تواند منجر به تصمیم اشتباه در انتخاب بهترین مدل از میان مدل‌های کاندید شود. در این خصوص پیشنهاد شد که آماره (d) باید همراه با آماره‌های RMSE و MBE برای ارزیابی عملکرد یک مدل استفاده شود و این آماره را می‌توان در کنار دو آماره دیگر برای قضاوت بهتر ارزیابی یک مدل بکار برد. در پژوهشی دیگر ماتسورا (۲۰۰۵) نشان داد که RMSE معادل MAE نیست و نمی‌توان به راحتی مقدار MAE را از RMSE استخراج نمود. در این مطالعه نشان داده شد که با وجود اینکه چندین مجموعه RMSE یکسانی دارند، MAE از مجموعه‌ای به مجموعه دیگر متفاوت می‌باشد. در تحقیقی وضعیت خشکسالی استان گلستان با استفاده از نمایه‌های خشکسالی انجام شد. برای این منظور بهترین روش درون‌یابی با معیار کمترین خطا (RMSE) انتخاب گردید (عیوضی و مساعدی، ۱۳۹۰). با توجه به افزایش حجم استفاده کاربران از آماره‌های ارزیابی خطا این پرسش مطرح است که آگاهی از نحوه عملکرد و بکارگیری صحیح از یک آماره‌ی اندازه‌گیری خطا چه کمکی برای پژوهشگر می‌تواند داشته باشد. همچنین انتخاب کدام یک از آماره‌ها می‌تواند منجر به تصمیمی صحیح در انتخاب بهترین مدل و یا روش قرار بگیرد. از این رو در این پژوهش هدف اصلی ارزیابی آماره‌های اندازه گیری خطا در انتخاب روش‌های درون‌یابی برای بارش استان فارس می‌باشد. از دیگر اهداف این مطالعه، استفاده از نتایج بدست آمده به عنوان الگویی برای انتخاب صحیح مدل‌های پایش مخاطرات محیطی می‌باشد.

داده‌ها و روش کار

استان فارس در جنوب ایران واقع شده و چهارمین استان بزرگ کشور به شمار می‌رود. آب و هوای این استان به سه اقلیم کوهستانی، معتدل و گرم تقسیم می‌گردد. میانگین بارش در این استان ۲۸۸ میلی‌متر می‌باشد. متوسط دمای این استان ۱۶/۶ درجه سلسیوس می‌باشد. از گرمترین شهرهای استان می‌توان مهر، لامرد، داراب و قیروکازرون را نام برد و از سردترین مناطق این استان شهرهای اقلید و سپیدان را نام برد. همچنین شهرهای مهرنجان ممسنی، فهلیان و بکر از شهرهای پربارش و شهرهای گردنه کولی‌کش و آباده از کم‌بارش ترین شهرهای استان فارس می‌باشند. در این پژوهش پس از بررسی وضعیت آماری

4-Root Mean Square Error
5-Mean Absolute Error
6-Mean Square Error
7-Mean Bias Error
1-Davies et al

ایستگاه‌ها، سال ۹۸ به عنوان سال منتخب و سالی پربارش انتخاب و از داده‌های بارش سالیانه ۱۳۹ ایستگاه باران‌سنجی و ۲۲ ایستگاه همدیدی استفاده شده است (سازمان هواشناسی کشور، ۱۴۰۰). در شکل (۱) محدوده مطالعاتی و همچنین موقعیت ایستگاه‌های همدید و باران‌سنجی ارائه شده است. از دلایل اصلی انتخاب استان فارس، توزیع جغرافیایی مناسب و تعداد بالای ایستگاه‌های هواشناسی و همچنین تنوع اقلیمی استان می‌باشد.



شکل ۱) موقعیت جغرافیایی ایستگاه‌های همدید و باران‌سنجی دراستان فارس

جهت استفاده از روش‌های درون‌یابی در این پژوهش از نرم افزار GIS استفاده گردیده است. هدف از اعمال روش‌های درون‌یابی، استخراج مکانی داده‌ها در قالب پهنه‌ای-رستری (شمسی‌پور، ۱۳۹۳). جهت انجام صحیح درون‌یابی و استفاده از داده‌های نقطه‌ای و تعمیم آن به سطحی رستری و پیوسته، از سه مرحله‌ی ۱- آماده سازی داده‌ها شامل پیدا کردن منابع خطا با استفاده از تجزیه و تحلیل آماری و بررسی توزیع و نرمال کردن داده‌ها ۲- اعمال روش‌های درون‌یابی اعم از روش‌های جبری و زمین‌آماری ۳- اعتبارسنجی روش‌ها در راستای انتخاب روشی مناسب استفاده شده است (Salah, 2009). برای اعمال درون‌یابی، روش‌های وزن‌دهی معکوس فاصله (IDW)[9]، کریجینگ (Kriging) ، کوکریجینگ (Co-Kriging) ، تابع پایه شعاعی (RBF)[10] روش‌های رایج استفاده شده در درون‌یابی می‌باشد که در تحقیقات پیشین در حوزه‌های متفاوت جغرافیایی استفاده شده‌اند. روش وزن دهی معکوس فاصله (IDW)، بر اساس قانون تیلور عمل می‌کند. این قانون که به عنوان قانون اول جغرافیا نیز شناخته می‌شود، اذعان دارد که واحدهای جغرافیایی نزدیک‌تر تأثیرگذاری بیشتری بر روی یکدیگر دارند. این روش بدون توجه به موقعیت و آرایش نقاط، فقط ترکیب خطی از معکوس فاصله داده‌های موجود را مد نظر قرار می‌دهد (Ding et. al., 2018). در روش RBF سطح ایجاد شده باید از میان تمامی نمونه‌های اندازه‌گیری شده، عبور کند. این روش

9. Inverse Distance Weighting
10. Radial Basis Functions

[Downloaded from system.khu.ac.ir on 2026-11-03]
[DOR: 20.1001.1.24237892.1403.11.2.7.9]
[DOI: 10.61186/jsaeh.11.2.117]

جزء دسته روش‌های درونیابی دقیق بوده و انواع مختلفی دارد (Esri, 2020). روش تخمین‌گر عام [۱۱](GP)، در یک منطقه وسیع و جهانی که در آن بررسی ها سطحی و وسیع هستندُ استفاده می‌گردد (Xin, 2009). همچنین روش تخمین‌گر موضعی [۱۲](LP)، فرایندی است که طی آن یک رابطه (غالبا چند جمله‌ای) برای درون‌یابی تعریف می‌شود و روشی سریع با دقتی کمتر و ایجاد کننده سطحی صاف می‌باشد. پیشنهاد می‌شود از این روش در مناطق کوچک و با نمونه‌های با اعضای کم استفاده گردد (Fan and Gibels, 1996). یکی از رایج‌ترین روش‌های زمین‌آماری، کریجینگ می‌باشد (Li et. al., 2020). این روش عموماً در تخمین و شباهت توزیع مکانی در متغیرهایی که بین آنها خودهمبستگی مکانی وجود دارد استفاده می‌شود (Belkhiri et. al., 2020). از آنجا که معمولا روش‌های درون‌یابی با مقادیر متفاوتی از خطا همراه است، عدم توجه به میزان خطای مدل‌های درون‌یابی، سبب انحراف نتایج از واقعیت‌ها می‌شود که این امر قابل قبول نمی‌باشد (Mitasova and Miats, 1999). از این رو جهت ارزیابی عملکرد و میزان معناداری آماره‌ها، تعدادی از آنها که شامل روش‌های میانگین خطای اریب [۱۳](MBE) میانگین مربعات خطا [۱۴](MSE)، ریشه میانگین مربعات خطا [۱۵](RMSE)، میانگین خطای مطلق [۱۶](MAE)، کارایی نش-ساتکلیف [۱۷](NSE)، ضریب تبیین [۱۸](r^2)، معیار درصد میانگین خطای پیش‌بینی [۱۹](MAPE)، ریشه میانگین مربعات خطای نرمال شده [۲۰](NRMSE)، می‌باشند، بررسی و ارزیابی می‌گردند. قابل ذکر است که ضریب تبیین توان دوم ضریب همبستگی پیرسون می‌باشد (رابطه ۸). روابط هر یک از این آماره‌ها در ادامه آورده شده که در این روابط (e_i) مقدار تخمینی (برآورد شده) و (o_i) مقدار مشاهده‌ای (واقعی) می‌باشد.

$$MBE = \frac{\sum_{i=1}^{n}(e_i - o_i)}{N}$$ رابطه (۱)

$$MSE = \frac{\sum_{i=1}^{n}(e_i - o_i)^2}{N}$$ رابطه (۲)

$$RMSE : \sqrt{\frac{\sum_{i=1}^{n}(e_i - o_i)^2}{N}}$$ رابطه (۳)

11. Global Polynomial
12 Local Polynomial
13-Mean Bias Error
14-Mean Square Error
15-Root Mean Square Error
16-Mean Absolute Error
17-Nash-Sutcliffe efficiency
18-Coefficient of Determination
19-Mean Absolute Percentage Error Loss
20-Normalized Root Mean Square Error

[Downloaded from system.khu.ac.ir on 2026-11-03]
[DOR: 20.1001.1.24237892.1403.11.2.7.9]
[DOI: 10.61186/jsaeh.11.2.117]

رابطه (۴)
$$MAE : \frac{\sum_{i=1}^{n} |e_i - o_i|}{N}$$

رابطه (۵)
$$MAPE = \frac{\sum_{i=1}^{n} \frac{(e_i - o_i)}{e_i} \times 100}{N}$$

رابطه (۶)
$$NSE = \frac{\sum_{i=1}^{n} (e_i - o_i)^2}{\sum_{i=1}^{n} (e_i - \bar{o})^2}$$

رابطه (۷)
$$NRMSE = \frac{RMSE}{e_i(\max) - e_i(\min)}$$

رابطه (۸)
$$r = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n} (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}}$$

دامنه آماره MBE بین مثبت و منفی بی‌نهایت می‌باشد که مقادیر مثبت و منفی به ترتیب بیانگر بیش‌برآوردی و کم‌برآوردی می‌باشد. همچنین آماره‌های MSE، RMSE و MAE دارای دامنه‌ی بین صفر تا مثبت بی‌نهایت می‌باشند. مقادیر هر چه به صفر نزدیک‌تر باشد از خطای کمتری در این آماره‌ها برخوردار است. آماره MAPE درصد خطا را نشان می‌دهد و مقادیر نزدیک به صفر دارای خطای کمتر می‌باشد. دامنه آماره NSE بین یک و منهای بی‌نهایت می‌باشد که اگر مقدار آن یک باشد برآورد مدل بالاترین نمره را دارد و بدون خطا می‌باشد و مقادیر بین صفر و یک بیانگر سطح قابل قبولی از دقت روش و یا مدل مورد نظر می‌باشد. مقادیر کمتر از صفر نیز غیر قابل قبول می‌باشد. در خصوص ضریب تعیین (r^2) دامنه تغییرات آن بین صفر تا مثبت یک بوده که هر چه مقدار به یک نزدیک‌تر باشد، نشان دهنده تخمین دقیق‌تری از روش درون‌یابی و کمتر بودن خطا می‌باشد.در نهایت شاخص توافق (d) دامنه بین ۰ و ۱ را دارا می‌باشد و هر چه مقدار آن به یک نزدیک‌تر باشد عملکرد مدل بهتر می‌باشد (Willmott, 1982) .

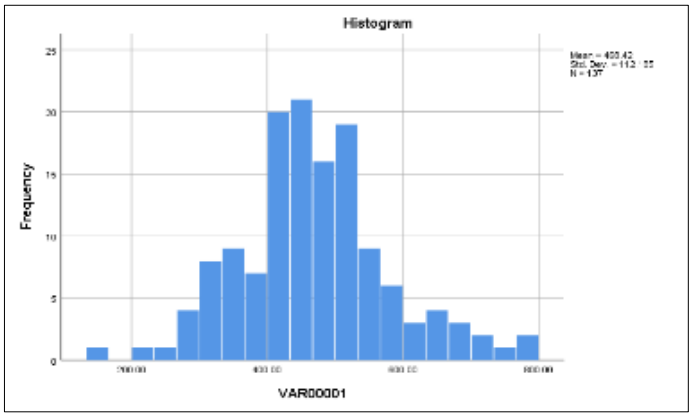
نتایج

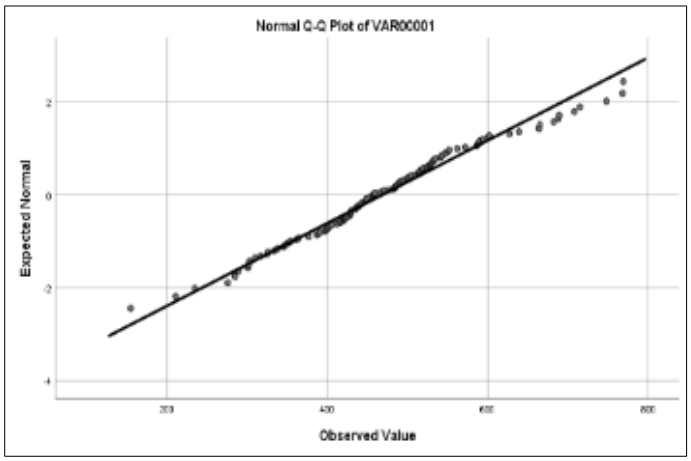
آزمون نرمال بودن داده ها

نرمال بودن توزیع داده ها به کمک آزمون کلموگروف اسمیرنوف[۲۱] (K-S) یک نمونه‌ای بررسی گردید. (حسنی پاک و شرف‌الدین، ۱۳۹۰). نتایج نشان داد (جدول شماره ۱) با توجه به مقدار P-value که از ۰/۰۵ بزرگتر می‌باشد، داده‌ها از توزیع نرمال برخوردار می‌باشند. در شکل (۲) وضعیت نرمال بودن داده‌های بارش به کمک هیستوگرام و نمودار چندک چندک نشان داده شده است.

21-One-Sample Kolmogorov-Smirnov test

[Downloaded from system.khu.ac.ir on 2026-11-03]
[DOR: 20.1001.1.24237892.1403.11.2.7.9]
[DOI: 10.61186/jsaeh.11.2.117]





شکل ۲) نمودار بافت نگار (چپ) و نمودار چندک چندک (راست)

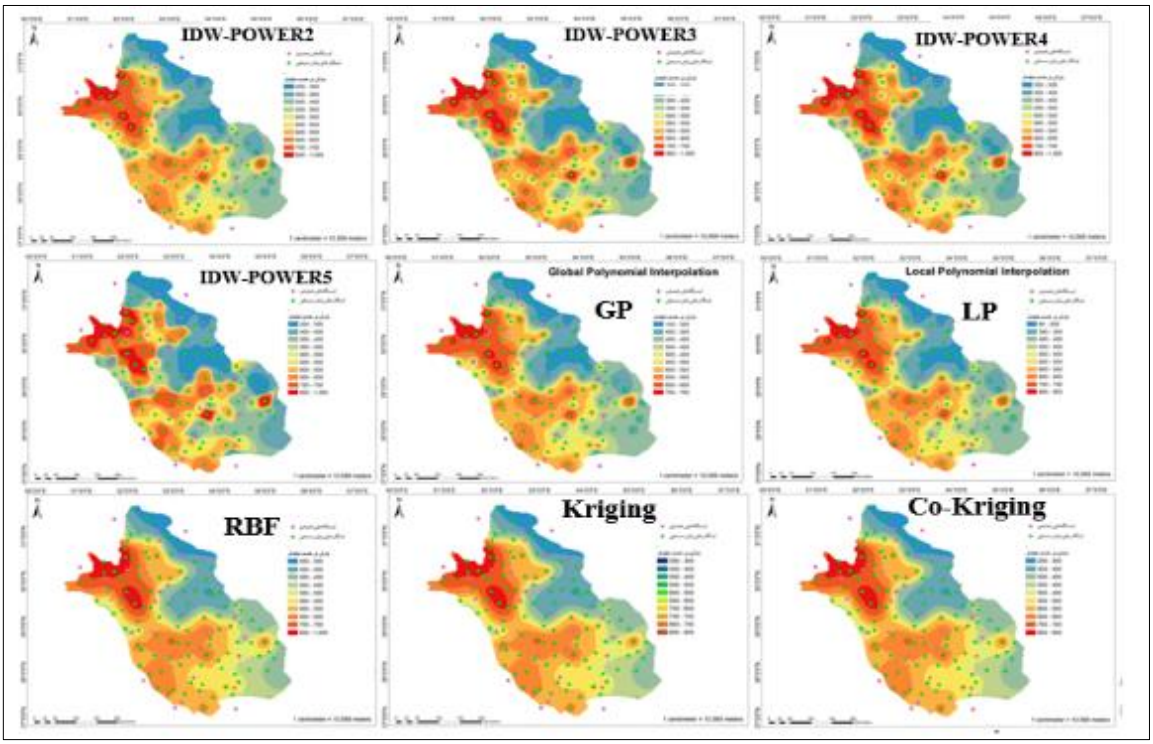
جدول ۱) پارامترهای آماری بارش سال ۱۳۹۸ استان فارس به همراه آماره کلموگروف-اسمیرنوف

پارامترهای آماری	تعداد نمونه	میانگین	انحراف معیار	چولگی	کشیدگی	P-value
مقادیر	۱۳۷	۴۷۷	۱۶۱	۰/۷۴۶	/۷۸۲.	۰/۱۰۷

اعمال روش‌های درون‌یابی

بعد از آماده سازی داده‌ها جهت اعمال روش‌های درون‌یابی از روش‌های جبری شامل معکوس وزنی فاصله با توان‌های (۲،۳،۴،۵)، توابع پایه شعاعی، چند جمله‌ای محلی و جهانی و روش‌های زمین آماری شامل کریجینگ و کوکریجینگ استفاده شده است. در شکل شماره (۳) مقادیر تخمینی و مقادیر مشاهده شده بارش با استفاده از روش اعتبار سنجی متقابل در تمامی مدل‌ها استخراج و مدل‌های درون‌یابی بر روی آنها اعمال گردیده است.

[Downloaded from system.khu.ac.ir on 2026-11-03]
[DOR: 20.1001.1.24237892.1403.11.2.7.9]
[DOI: 10.61186/jsaeh.11.2.117]

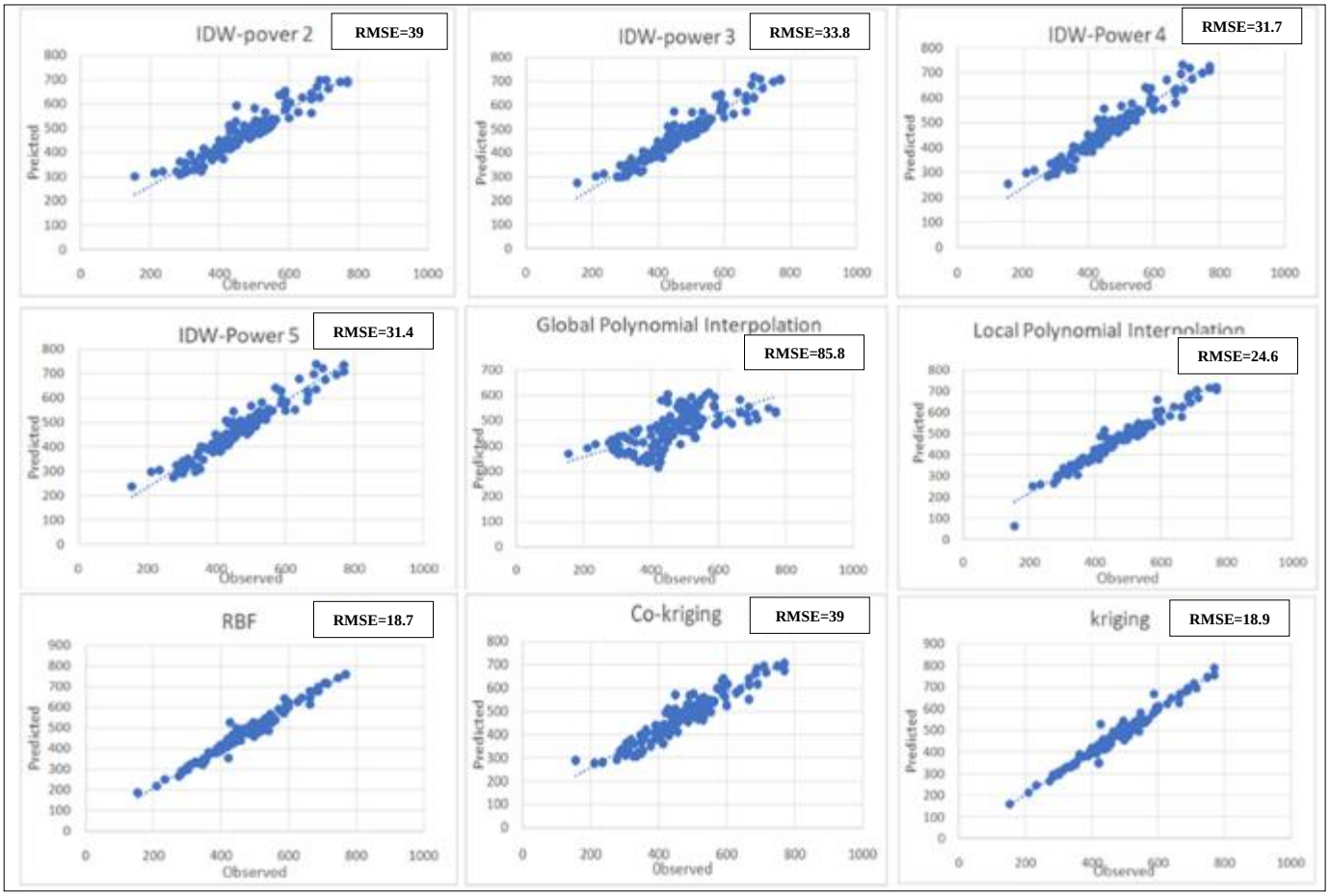


شکل ۳) نقشه پهنه‌بندی بارش سال ۱۳۹۸ استان فارس با استفاده از روش‌های درون‌یابی

ارزیابی دقت روش‌های درون‌یابی

با توجه به نمودارهای شکل شماره (۴) و یک بررسی اجمالی می‌توان به این نتیجه رسید که روش‌های Kriging (میلی‌متر RMSE = ۱۸/۹)، RBF (میلی‌متر RMSE = ۱۸/۷) و Local (میلی‌متر RMSE = ۲۴/۶) مقدار بارش را با دقت بالاتری تخمین زده‌اند. این در حالی است که روش‌های Global Polynomial Interpolation (میلی‌متر RMSE = ۸۵/۸) و Co-kriging (میلی‌متر RMSE = ۳۹/۰) به ترتیب از دقت کمتری برخوردار بوده‌اند. در خصوص روش IDW، در توان‌های بالاتر، درون‌یابی مقادیر بارش با خطای کمتری انجام می‌شود (شکل شماره ۴).

[Downloaded from system.khu.ac.ir on 2026-11-03]
[DOR: 20.1001.1.24237892.1403.11.2.7.9]
[DOI: 10.61186/jsaeh.11.2.117]



شکل شماره۴) نمودار مقایسه مقادیر مشاهده شده و تخمین‌زده شده بارش سال ۱۳۹۸ به کمک روش‌های مختلف درون‌یابی

سنجش دقت روش‌های درون‌یابی توسط آماره‌های ارزیابی خطا

در مقایسه نتایج بدست آمده توسط آماره‌های ارزیابی خطا، بعد از مشاهده میزان خطا و نمره‌دهی هر یک از این آماره‌ها، به ارزیابی عملکرد و مقایسه‌آماری معیارهای ارزیابی خطا با یکدیگر پرداخته شد که نتایج آنها در جدول شماره (۲) آورده شده است. در ادامه ویژگی‌های هر یک از آماره‌ها به تفکیک مورد بررسی قرار گرفته است.

جدول ۲) مقادیر آماره‌های اندازه‌گیری خطای روش‌های درون‌یابی بارش سال ۱۳۹۸ استان فارس

MAE	MBE	MSE	RMSE	NRMSE	MAPE	SMAPE	NSE	d	r^2	روش
۲۷	-۶/۴	۱۵۲۰	۳۹	۰/۰۸۴	۶/۷	۰/۰۶۳	۰/۸۷۹	۰/۷۱۷	۰/۹۴۶	IDW2
۲۹	-۵/۱	۱۱۴۵	۳۳/۸	۰/۰۴۸	۵/۷۴	۰/۰۵۴	۰/۹۰۷	۰/۰۵۲	۰/۹۵۷	IDW3
۲۳	-۴/۳	۱۰۰۸	۳۱/۷	۰/۰۴۶	۵/۴۸	۰/۰۵۲	۰/۹۱۸	۰/۰۴۴	۰/۹۶۱	IDW4
۲۳	-۳/۷	۹۸۹	۳۱/۴	۰/۰۴۵	۵/۵	۰/۰۵۳	۰/۹۲۰	۰/۴۲۹	۰/۹۶۷	IDW5
۱۲	-۱/۱	۳۴۸	۱۸/۷	۰/۰۴۲	۲/۸۴	۰/۰۲۸	۰/۹۷۲	۰/۱۴۲	۰/۹۸۶	RBF

[Downloaded from system.khu.ac.ir on 2026-11-03]
[DOR: 20.1001.1.24237892.1403.11.2.7.9]
[DOI: 10.61186/jsaeh.11.2.117]

Global	۰/۶۳۷	۰/۰۴۱۳	۰/۴۰۶	۰/۱۴۸	۱۵/۷	۰/۰۵۸	۵۸/۸	۷۳۵۷	-۰/۱	۶۷
Local	۰/۹۷۶	۰/۰۲۵	۰/۹۵۱	۰/۰۴۱	۳/۹۴	۰/۰۳۹	۲۴/۶	۶۰۶	۹/۲	۱۷
Kriging	۰/۹۸۶	۰/۱۴۹	۰/۹۷۱	۰/۰۲۵	۲/۴۹	۰/۰۴۰	۱۸/۹	۳۵۷	۰/۳	۱۲
Co-kriging	۰/۹۴۲	۰/۰۷۰۹	۰/۸۷۸	۰/۰۶۸	۷/۱۲	۰/۰۴۸	۳۹	۱۵۲۰	-۴/۳	۳۰

آماره MBE

بررسی نتایج حاصل از آماره MBE نشان می‌دهد به دلیل ماهیت ریاضی این آماره در مواردی به دلیل صفر شدن حاصل جمع مقادیر مثبت (مقادیر بیش از مقدار واقعی، بیش‌برآوردی) و منفی (مقادیر کمتر از مقدار واقعی کم‌برآوردی)، محقق در انتخاب روش درون‌یابی دچار اشتباه شده و روش درون‌یابی با خطای بیشتر را انتخاب می‌کند. برای نمونه در روش کریجینگ با وجود اینکه مقدار نهایی MBE، ۰/۳ و به صفر بسیار نزدیک است، اختلاف مقادیر واقعی از تخمینی بارش در مواردی ۲۲ و حتی ۳۲- (منهای ۳۲) شده است (جدول شماره ۳). بنابراین نمی‌توان از این آماره برای قضاوت در دقت و انتخاب روش‌های درون‌یابی استفاده نمود.

جدول ۳) مقادیر مشاهده شده و تخمینی بارش (بر حسب میلی‌متر) در روش کریجینگ (MBE=۰/۳)

مشاهده شده	۳۸۹	۴۹۳	۴۸۹	۴۵۷	۴۶۶	۶۶۴	۷۶۳	۵۷۲	۵۵۱	۵۳۳	۴۰۱	۲۸۹	۳۷۷	۳۵۱
تخمین زده شده	۳۸۲	۵۰۶	۴۷۱	۴۵۶	۴۹۴	۶۲۸	۷۵۳	۵۵۰	۵۵۶	۵۲۵	۳۹۲	۲۹۰	۳۸۴	۳۶۹

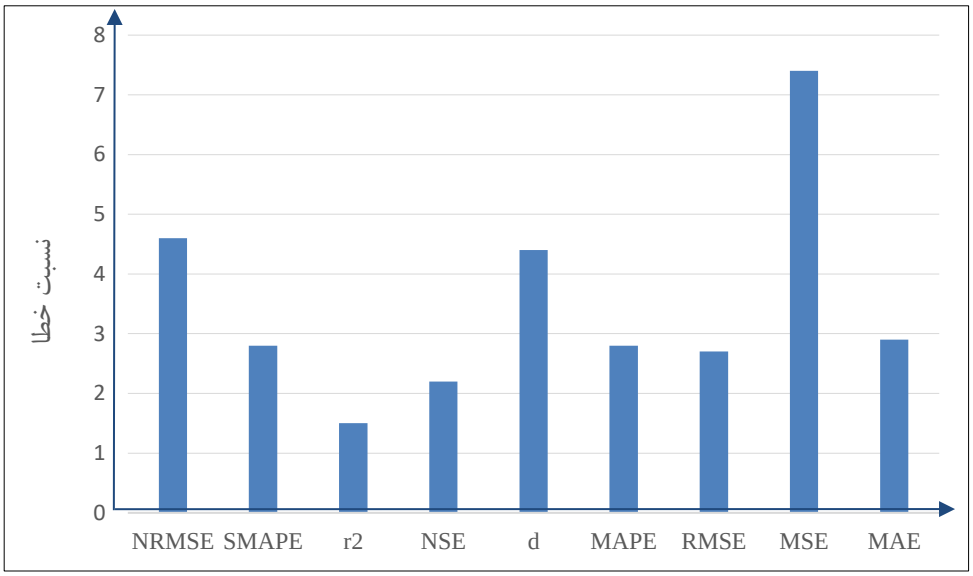
آماره MSE

در بررسی آماره MSE (جدول شماره ۴)، خطای دو روش چندجمله‌ای جهانی و روش معکوس وزنی فاصله با یکدیگر مقایسه شده است. مشاهده می‌شود که مقدار آماره MSE در روش معکوس وزنی فاصله با توان پنجم، ۹۸۲ می‌باشد و در روش چند جمله‌ای جهانی ۷۳۵۷ بوده و نسبت خطای این دو روش ۷/۴ می‌باشد (جدول شماره ۴). این در حالی است که نسبت این خطا در دیگر آماره‌ها بسیار پایین‌تر می‌باشد (بین ۱/۵ تا ۴/۶). با توجه به ماهیت رابطه MSE در مواردی که داده‌های پرت وجود دارد، به دلیل مربع شدن خطا نتیجه نهایی بزرگتر می‌شود (Willmott and Matsuura, 2005). بنابراین در انتخاب روش درون‌یابی به کمک MSE بایستی دقت بیشتری نمود.

جدول شماره ۴) مقایسه نسبت خطای روش معکوس وزنی فاصله با توان پنجم و روش چندجمله‌ای جهانی در MSE با دیگر آماره‌ها

روش	NRMSE	SMAPE	r^2	NSE	D	MAPE	RMSE	MSE	MAE
IDW5	۰/۰۶۲	۰/۰۵۳	۰/۹۶۱	۰/۹۲۰	۰/۰۹۲	۵/۵	۳۱/۴	۹۸۲	۲۳
Global	۰/۲۸۷	۰/۱۴۸	۰/۶۳۸	۰/۴۰۶	۱/۱۳	۱۵/۷	۸۵/۸	۷۳۵۳	۶۷
نسبت خطا	۴/۶	۲/۸	۱/۵	۲/۲	۴/۴	۲/۸	۲/۷	۷/۴	۲/۹

[Downloaded from system.khu.ac.ir on 2026-11-03]
[DOR: 20.1001.1.24237892.1403.11.2.7.9]
[DOI: 10.61186/jsaeh.11.2.117]



شکل ۵) نسبت مقادیر آماره‌های ارزیابی خطا در دو روش معکوس وزنی فاصله و چند جمله‌ای جهانی

آماره‌های (MSE) و (RMSE)

آماره RMSE حاصل جذر MSE می‌باشد (رابطه ۳). بنابراین پس از جذر گرفتن از مقادیر MSE حاصل به عدد کوچکتری تغییر یافته که همان مقدار RMSE خواهد بود (جدول ۵). تفاوت دیگر این دو آماره در این است که در محاسبه (RMSE) به دلیل مجذور شدن عدد (MSE)، حاصل را به واحد متغیر نشان داده و مقایسه منطقی‌تری در تحلیل نتایج می‌توان از آن داشت.

جدول ۵) مقایسه مقادیر MSE و RMSE در روش‌های مختلف درون‌یابی

روش	RMSE	MSE
IDW2	۳۹	۱۵۲۰
IDW3	۳۳/۸	۱۱۴۵
IDW4	۳۱/۷	۱۰۰۸
IDW5	۳۱/۴	۹۸۹
RBF	۱۸/۷	۳۴۸
global	۸۵/۸	۷۳۵۷
local	۲۴/۶	۶۰۶
kriging	۱۸/۹	۳۵۷
Co-kriging	۳۹/۰	۱۵۲۰

آماره MAE

آماره MAE معیار دیگری است که به طور گسترده در ارزیابی دقت روش‌ها استفاده می‌شود. دامنه مقادیر حاصل از این آماره که میانگین مجموع قدر مطلق تفاوت مقادیر مشاهده شده و تخمینی می‌باشد از صفر تا مثبت بی‌نهایت می‌باشد و حاصل آن

[Downloaded from system.khu.ac.ir on 2026-11-03]
[DOR: 20.1001.1.24237892.1403.11.2.7.9]
[DOI: 10.61186/jsaeh.11.2.117]

هر چه به صفر نزدیک‌تر باشد بیانگر خطای کمتر روش درون‌یابی است. با توجه به جدول شماره (۶) در مقایسه با آماره RMSE، آماره MAE دارای مقادیر خطای کوچک‌تر و حداکثر برابر آن می‌باشد. از ویژگی‌های هر دو آماره بیان خطا با مقیاس متغیر بوده، که از محاسن این دو در تحلیل نتایج می‌باشد.

جدول ۶) مقایسه مقادیر خطا به کمک آماره‌های MAE و RMSE

روش	MAE	RMSE
IDW2	۲۷	۳۹
IDW3	۲۹	۳۳/۸
IDW4	۲۳	۳۳/۷
IDW5	۲۳	۳۳/۴
RBF	۱۲	۱۸/۷
Global	۶۷	۸۵/۷
Local	۱۷	۲۴/۶
Kriging	۱۲	۱۸/۹
Co-Kriging	۳۰	۳۹/۰

آماره NRMSE

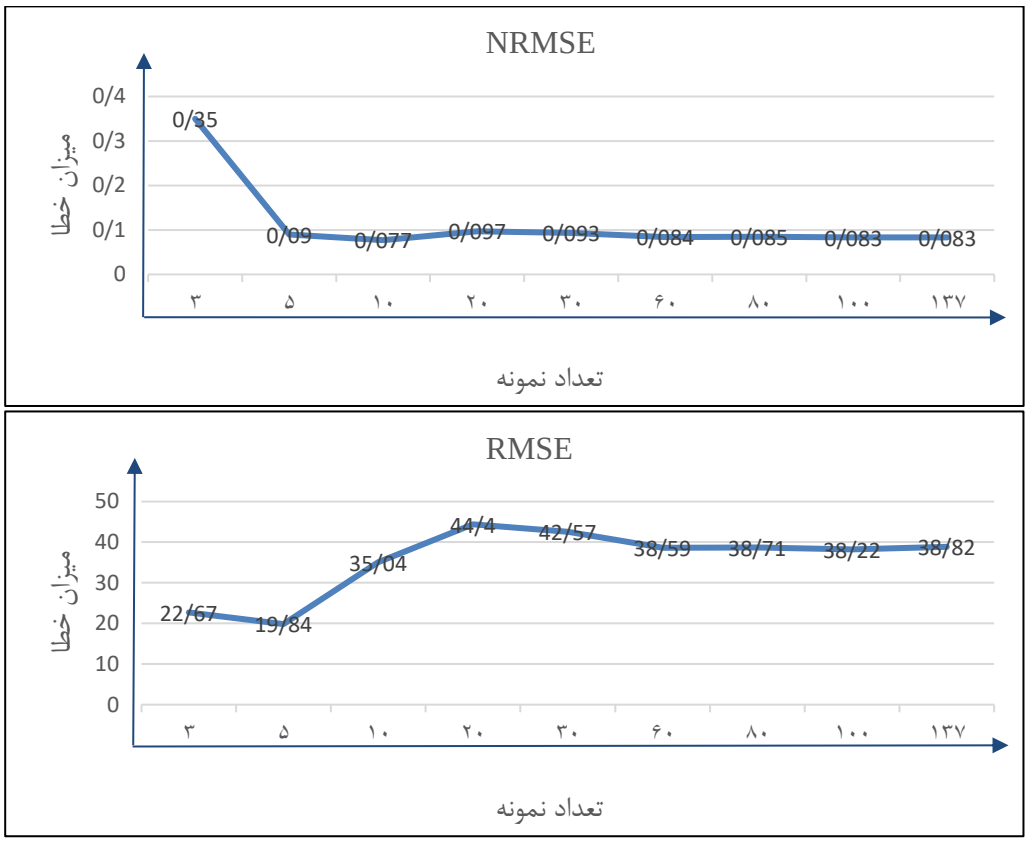
NRMSE نرمال شده آماره‌ی RMSE است که دامنه تغییرات آن بین صفر و یک می‌باشد. این آماره با هدف سنجش خطای مدل به کار می‌رود و به دلیل بی‌بعد کردن خطا، مقایسه دو متغیر با مقیاس‌های متفاوت از نکات مثبت استفاده از این آماره می‌باشد. در جدول شماره (۷) و شکل شماره (۶) ارزیابی خطا توسط آماره‌های RMSE و NRMSE در تعداد نمونه های متفاوت محاسبه شده است. همانگونه که مشاهده می‌شود (جدول ۷ و شکل ۶)، میزان خطا در روش‌هایی که از تعداد افراد نمونه کمی برخوردار می‌باشند، به شکل فزاینده‌ای افزایش پیدا کرده و استفاده از این آماره (NRMSE) برای نمونه‌های با تعداد کم (نمونه‌هایی با حجم کمتر از ۵ عضو) پیشنهاد نمی‌شود. نتایج این مطالعه نشان داد (جدول شماره ۲) که آماره NRMSE، روش IDW با توان دوم را به عنوان روش با دقت پایین‌تر در نظر گرفته، در صورتی که سایر آماره‌ها، روش چندجمله‌ای جهانی را به عنوان ضعیف‌ترین روش انتخاب کرده اند. بنابراین استفاده از این آماره نتایج قابل اعتمادی در بر نخواهد داشت.

جدول ۷) ارزیابی خطا به کمک NRMSE و RMSE در نمونه‌های با تعداد افراد متفاوت

تعداد افراد نمونه	RMSE	NRMSE
۳	۲۲/۶۷	۰/۳۵
۵	۱۹/۸۴	۰/۰۹۰
۱۰	۳۵/۰۴	۰/۰۷۷
۲۰	۴۴/۴۰	۰/۰۹۷
۴۰	۴۲/۵۷	۰/۰۹۳

[Downloaded from system.khu.ac.ir on 2026-11-03]
[DOR: 20.1001.1.24237892.1403.11.2.7.9]
[DOI: 10.61186/jsaeh.11.2.117]

۶۰	۳۸/۵۹	۰/۰۸۴
۸۰	۳۸/۷۱	۰/۰۸۵
۱۰۰	۳۸/۲۲	۰/۰۸۳
۱۳۷	۳۸/۸۲	۰/۰۸۳



شکل شماره ۶) ارزیابی خطا به کمک NRMSE و RMSE در نمونه‌های با تعداد افراد متفاوت

آماره MAPE

آماره MAPE که درصد میانگین خطای MAE می‌باشد به دلیل اینکه حاصل آن به درصد بیان می‌شود، قادر به مقایسه خطای دو متغیر با واحدهای متفاوت می‌باشد. نتایج این مطالعه نشان داد (جدول شماره ۸) در مواقعی که مقادیر مشاهده شده به صفر نزدیک می‌شوند مقدار خطا زیاد شده و ممکن است MAPE بیش از اندازه افزایش یابد (۷۰= MAPE در مقایسه با RMSE = ۰/۳۳ و MAE = ۰/۱۸).

[Downloaded from system.khu.ac.ir on 2026-11-03]
[DOR: 20.1001.1.24237892.1403.11.2.7.9]
[DOI: 10.61186/jsaeh.11.2.117]

جدول ۸) جدول ارزیابی خطا برای مقادیر مشاهده شده‌ی نزدیک به صفر

مشاهده‌شده	تخمینی
۰/۰۴۰	۰/۰۲۰۰
۰/۰۰۸	۰/۰۱۰۰
۰/۹۰۰	۰/۰۲۰۰
۰/۳۱۰	۰/۰۲۰۰
۰/۱۲۰	۰/۰۲۰۰
۰/۰۹۰	۰/۰۰۱۰
۰/۰۴۰	۰/۰۰۲۰
۰/۰۶۰	۰/۰۰۴۰

خطا

MAE	MAPE	RMSE
۰/۱۸۴۴	۷۰/۰۱۳	۰/۳۳۲

از دیگر نقاط ضعف آماره MAPE در زمانی است که اگر مقادیر تخمینی کمتر از مشاهده شده باشد (کم‌برآوردی) خطای بیشتری را نسبت به زمانی که مقادیر تخمینی بیشتر از مقادیر مشاهده شده باشد (بیش‌برآوردی)، نشان می‌دهد. این در حالی است که آماره MAE در هر دو حالت یک مقدار خطا را نشان می‌دهد (جدول شماره ۹). همچنین در زمانی که مقادیر تخمینی کمتر از مقادیر مشاهده شده است، میزان خطا ۶/۵ بوده و در حالت عکس مقدار خطا کمتر می‌باشد (۶/۰). در نتیجه در تخمین‌های کمتر روش درون‌یابی، آماره MAPE خطای بیشتری را نشان می‌دهد.

جدول ۹) جدول ارزیابی خطا توسط مقادیر بیش برآوردی و بالعکس

مشاهده‌شده	تخمینی
۱۱۲	۱۲۱
۱۳۲	۱۳۹
۱۲۴	۱۲۹
۱۳۰	۱۴۱
MAPE	
۶/۰۳۸	

تخمینی	مشاهده‌شده
۱۲۱	۱۱۲
۱۳۹	۱۳۲
۱۲۹	۱۲۴
۱۴۱	۱۳۰
MAPE	
۶/۴۵۸	

MAE
۸

[Downloaded from system.khu.ac.ir on 2026-11-03]
[DOR: 20.1001.1.24237892.1403.11.2.7.9]
[DOI: 10.61186/jsaeh.11.2.117]

ضریب تعیین (r^2)

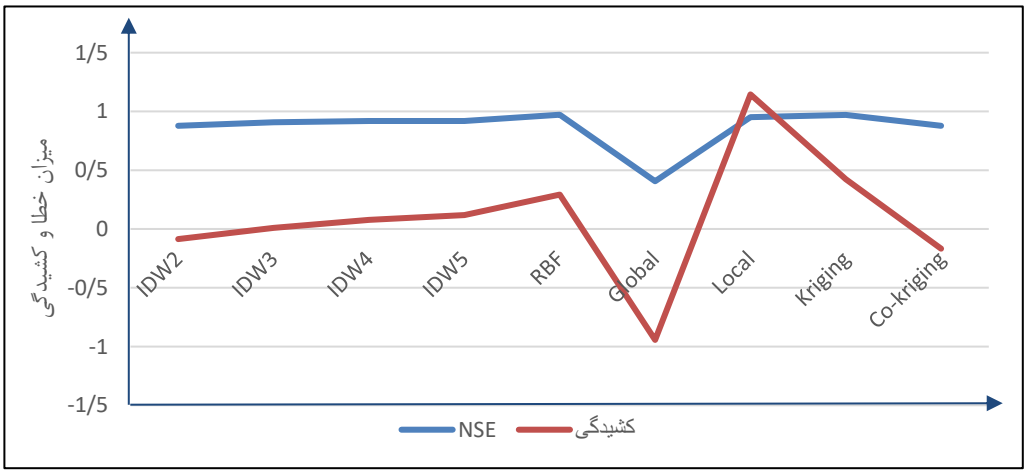
نتایج این مطالعه بیانگر مشابه بودن مقدار r^2 در دو روش RBF و IDW با توان ۵ می‌باشد. این در حالی است که دو آماره MAE و RMSE در روش IDW با توان ۵ مقدار خطای دو برابری نسبت به روش RBF را نشان می‌دهند (جدول شماره ۱۰). از طرف دیگر بررسی نتایج ارزیابی روش‌های درون‌یابی نشان داد که استفاده از ضریب تبیین (r^2)، به دلیل عدم سنخیت بزرگی دامنه تغییرات آن (صفر تا ۱) با مقادیر خطای درون‌یابی بارش استان فارس (۱۰۰ تا ۴۰۰ میلی‌متر)، در ارزیابی دقت روش‌های درون‌یابی منطقی نمی‌باشد.

جدول ۱۰) مقایسه نتایج آماره r^2 با دو آماره MAE و RMSE

روش	IDW2	IDW3	IDW4	IDW5	RBF	Global	Local	Kriging	Co-kriging
r^2	۰/۹۴۶	/۹۶۱	۰/۹۶۱	۰/۹۶۷	۰/۹۸۶	۰/۶۳۸	۰/۹۷۶	۰/۹۸۶	۰/۹۴۲
MAE	۲۷	۲۹	۲۳	۲۳	۱۲	۶۷	۱۷	۱۲	۳۰
RMSE	۳۹	۳۳/۸	۳۱/۷	۳۱/۴	۱۸/۷	۸۵/۵	۲۴/۶	۱۸/۹	۳۹

شاخص ضریب کارایی نش ساتکلیف (NSE)

با توجه به شکل شماره (۷) و مشاهده نمودار میزان کشیدگی و ارتباط مستقیم آن با خطای NSE می‌توان گفت این آماره نسبت به مقادیر پرت از حساسیت بالایی برخوردار بوده و می‌توان با لگاریتم‌گیری و نرمال کردن داده‌ها آن را رفع نمود. آماره NSE، در واقع نمونه تعدیل یافته و توسعه یافته MSE می‌باشد (زند کریمی و مختاری، ۱۳۹۷). نتایج این تحقیق نشان داد که بر خلاف سایر آماره‌های دیگر که برخی روش‌ها را با دقت بالا و برخی را با دقت پایین‌تر ارزیابی نمودند، آماره NSE بیشتر روش‌ها را با دقت بالا ارزیابی و همه روش‌های درون‌یابی در این مطالعه را تایید نمودند (NSE > ۰/۹۰).



شکل شماره ۷) نمودار خطی میزان خطای NSE و میزان کشیدگی روش‌های درون‌یابی

[Downloaded from system.khu.ac.ir on 2026-11-03]
[DOR: 20.1001.1.24237892.1403.11.2.7.9]
[DOI: 10.61186/jsaeh.11.2.117]

بحث

فکته وهمکاران[۲۲] (۲۰۰۴) و ژاکوویدس[۲۳] (۱۹۹۴) در مطالعه‌ای نشان دادند که آماره MBE اطلاعات مفیدی در خصوص عملکرد یک مدل ارائه می دهد. از دیگر نتایج این مطالعه می‌توان به این موضوع اشاره کرد که مقادیر مثبت و منفی از این آماره به ترتیب بیانگر تخمین بیش‌برآوردی و کم‌برآوردی روش درون‌یابی می‌باشد. ایشان همچنین نشان دادند که مقادیر نزدیک به صفر نشان دهنده عملکرد مناسب روش درون‌یابی است. در مقایسه مطالعه ایشان با مطالعه حاضر دیده می‌شود که آماره MBE نمی‌تواند معیار مناسبی برای بررسی دقت یک مدل آماری در تخمین داده‌ها باشد. در تحقیق حاضر نشان داده شد که اگر چه احتمال دارد با توجه به ماهیت رابطه ریاضی این آماره (رابطه ۱) و صفر شدن حاصل جمع مقادیر مثبت و منفی، مقدار MBE نزدیک به صفر و حتی صفر شود، اما این نتیجه (MBE=۰) ، سبب گمراه شدن محقق در انتخاب روش درون‌یابی می‌گردد. لذا این آماره را نمی‌توان به عنوان یک معیار برای ارزیابی خطای یک روش درون‌یابی استفاده نمود. در خصوص آماره ریشه میانگین مربعات خطا (RMSE)، سال‌های متمادی است که به عنوان یکی از معیارهای قابل قبول برای ارزیابی خطای مدل توسط بسیاری از محققان استفاده می‌شود (مک‌کین و همکاران[۲۴] (۲۰۰۵) و چای و همکاران (۲۰۱۳)). از طرف دیگر نتایج مطالعات ویلموت و ماتسورا (۲۰۰۵) نشان می‌دهد که این آماره نسبت به خطاهای بالا حساسیت بیشتری داشته و خطا را بیشتر از مقدار واقعی آن نشان می‌دهد که هم‌راستا با مطالعه اخیر (بارش استان فارس) می‌باشد.

البته تحلیل مقایسه‌ای نتایج RMSE و MSE برای درون‌یابی بارش استان فارس نشان داد که به دلیل ماهیت رابطه RMSE و تعدیل نمودن خطا، استفاده از این آماره نسبت به MSE ارجحیت دارد. نتایج این مطالعه در رابطه با آماره MAE و مقایسه آن با آماره MBE بیانگر واقعی‌تر بودن مقادیر MAE نسبت به MBE است. همان‌گونه که قبلا اشاره گردید، مقادیر کوچک و حتی نزدیک به صفر میانگین خطای اریب (MBE) می‌تواند محقق را در انتخاب روش مناسب درون‌یابی به خطا و تصمیم‌گیری اشتباه بیندازد. این در حالی است که وجود علامت قدر مطلق در رابطه میانگین خطای مطلق (MAE) این اجازه را نمی‌دهد که حاصل به سمت صفر میل کرده و از طرفی مقایسه روش‌های مختلف را با یکدیگر منطقی‌تر می‌سازد. با توجه به نتایج تحقیق حاضر، در دو روش با MAE مشابه و نزدیک به یکدیگر، RMSE می‌تواند از روشی به روش دیگر متفاوت باشد. بر اساس نتایج بدست آمده (جدول شماره ۶)، مقادیر MAE در روش‌های درون‌یابی IDW با توان‌های ۲ و ۳ نزدیک به یکدیگر و به ترتیب ۲۷ و ۲۹ میلی‌متر بوده در حالیکه این مقادیر برای RMSE از یکدیگر اختلاف داشته و به ترتیب ۳۹ و ۳۳/۸ میلی‌متر بدست آمده است. فکت و همکاران[۲۵] (۲۰۰۴)، نشان دادند که مقادیر RMSE هرگز نمی‌تواند از MAE کوچک‌تر بوده و تفاوت‌های MAE و RMSE به دلیل تأثیر واریانس‌های خطا می‌باشد. رایر[۲۶] (۲۰۰۷)، اگر چه در تحقیق خود سهولت در تفسیر نتایج MAPE را مطرح نموده اما بیان داشتند این آماره تحت تاثیر مقادیر پرت قرار گرفته و در اغلب موارد خطای محاسبه شده را بیش از آنچه که واقعیت دارد، نشان می‌دهد. در این خصوص نتایج مطالعه حاضر نشان داد که در زمانی

22- Fekete, et al
23-Jacobides
24 -McCain et al
25- Fact
26 -Rayer

[Downloaded from system.khu.ac.ir on 2026-11-03]
[DOR: 20.1001.1.24237892.1403.11.2.7.9]
[DOI: 10.61186/jsaeh.11.2.117]

که مقادیر تخمینی کمتر از مقادیر مشاهده شده است، میزان خطا ۶/۵ بوده و در حالت عکس مقدار خطا کمتر می‌باشد (۶/۰). در نتیجه در روش‌های درون‌یابی با تخمین‌های کمتر، آماره MAPE خطای بیشتری را نشان می‌دهد (جدول شماره ۹). در خصوص ضریب تبیین (r^2) و استفاده از آن در ارزیابی روش‌های درون‌یابی، نتایج تحقیقات انجام شده (مک لارن لیمیت و همکاران (۱۹۸۰)، ویلموت (۱۹۸۱) و ریوز (۲۰۲۱)) نشان می‌دهد مقادیر بالای r^2 به دلیل اینکه به بزرگی تفاوت مقادیر مشاهده شده و تخمینی توجهی ندارد، نمی‌تواند معیار مناسبی برای ارزیابی روش‌های درون‌یابی باشد. در مطالعه اخیر نتایج نشان داد که ارزیابی ۹ روش درون‌یابی به کمک آماره r^2 بیانگر مقادیر بالای آن و تایید اغلب روش‌ها توسط آن می‌باشد. در حالی که ارزیابی این روش‌ها به کمک سایر آماره‌ها بیانگر تمایز آن‌ها بوده و برخی نسبت به دیگری ارجحیت داشته‌اند. بنابراین نمی‌توان در ارزیابی روش‌های درون‌یابی به اندازه و بزرگی r^2 خوش‌بین بود و استفاده از این آماره توصیه نمی‌شود. ویلموت (۲۰۰۵) استفاده از شاخص توافق (d) که یک معیار توصیفی و نسبی می‌باشد را پیشنهاد کرده است. بر اساس نتایج تحقیق اخیر (جدول شماره ۲) کلیه آماره‌های ارزیابی خطا به غیر از ضریب توافق، به دقت روش درون‌یابی چند جمله‌ای محلی (local) نمره قابل قبولی داده‌اند. از طرف دیگر ضریب توافق به دقت روش IDW با توان ۲ بالاترین نمره را داده (d=۰/۷۱۷) در حالی که ارزیابی این روش توسط سایر آماره‌ها، از متوسط به پایین بوده است (جدول شماره ۲). بنابراین بر خلاف نظر ویلموت این آماره نمی‌تواند برای ارزیابی دقت روش‌های درون‌یابی پیشنهاد گردد.

نتایج تحقیق ملسه و ننگیا[۲۷] (۲۰۰۵) نشان داد برای مقایسه دو مدل با مقیاس (واحد) متفاوت، استفاده از جذر میانگین مربع خطای نرمال شده (NRMSE) می‌تواند آماره مناسبی برای انتخاب روش درون‌یابی با دقت بالاتر باشد. این در حالی است که در تحقیق حاضر نتایج نشان می‌دهد که برای مقایسه خطای روش‌های درون‌یابی در استفاده از NRMSE بهتر است احتیاط بیشتری نمود. به دلیل وجود دامنه تغییرات خطاها در مخرج کسر NRMSE اگر تعداد افراد نمونه زیاد باشد (جدول ۷) مقدار خطا به شدت کم شده که گمراه کننده است (NRMSE=۰/۰۸۵ برای n=۸۰). این در حالی است که در نمونه‌های با تعداد افراد کم، مقدار خطا زیاد شده که در این حالت نیز نتیجه و قضاوت نهایی برای انتخاب روش درون‌یابی گمراه کننده است (NRMSE=۰/۳۵ برای n=۳). در خصوص کارایی نش و ساتکلیف (NSE) نتایج این تحقیق نشان می‌دهد این آماره به مقادیر پرت حساسیت بالایی داشته و مقادیر خطا را بیش از مقدار واقعی نشان داده است. بنابراین ضمن دقت بیشتر در استفاده از این آماره، بایستی کارایی آن در مقایسه با دیگر آماره‌ها مقایسه شود.

نتیجه گیری

سنجش دقت بخش مهمی از فرآیند ارزیابی مدل‌های آماری برای توصیف عملکرد آنها می‌باشد. مفهوم اساسی ارزیابی دقت یک مدل، بر اساس آماره‌های معینی می‌باشد و مقدار خطاها نشان می‌دهند که مدل چه میزان در پیش‌بینی و یا تخمین یک متغیر خطا داشته است. روش‌های زیادی برای تعیین میزان خطای مدل‌های آماری جهت برآورد میزان خطا، پیشنهاد شده‌اند، اما هیچ یک از آنها عاری از محدودیت نمی‌باشند و بسیاری در خصوصیاتی خاص مبهم می‌باشند. هدف از مطالعه اخیر این نبوده که یک آماره به عنوان یک معیار شاخص و برتر برای پژوهش‌های آینده دیکته شود، بلکه هدف ارائه رهنمودها و ارزیابی

27-Melesse and Nangia

[Downloaded from system.khu.ac.ir on 2026-11-03]
[DOR: 20.1001.1.24237892.1403.11.2.7.9]
[DOI: 10.61186/jsaeh.11.2.117]

صحیح از آماره‌های اندازه گیری خطا و تفسیر عملکرد آنها بوده تا بتوان گام مفیدی در راستای پذیرش و یا رد یک آماره بر اساس محاسن و محدودیت‌های آن برداشت. از نتایج این تحقیق می‌توان به عدم کارایی آماره MBE در انتخاب دقت یک روش درون‌یابی اشاره نمود. در برخی موارد در محاسبات این آماره جمع مقادیر مثبت و منفی خطا به گونه‌ای خواهند بود که نتیجه نهایی صفر شده و ما را در انتخاب مدل مناسب گمراه می‌کند. در خصوص آماره MSE نتایج این تحقیق نشان داد که زمانی که تفاوت مقادیر تخمینی روش درون‌یابی نسبت به مقادیر مشاهده شده بزرگتر باشد، مقادیر این آماره بزرگتر شده (به دلیل ماهیت رابطه و وجود توان دوم در آن) و نمی‌تواند به عنوان یک آماره مطمئن مورد استفاده قرار گیرد. در مقابل با جذر گرفتن از این آماره که حاصل آن RMSE خواهد بود ($RMSE=\sqrt{MSE}$)، مقدار خطا به مقدار واقعی نزدیکتر می‌شود. همچنین در هنگام استفاده از آماره NSE، که نمونه تعدیل یافته و توسعه یافته MSE می‌باشد به نرمال بودن داده‌ها توجه گردد تا از میزان حساسیت به مقادیر شدید در این آماره کاسته شود. از دیگر نتایج این مطالعه می‌توان به حساسیت NRMSE در نمونه‌هایی با تعداد کم اشاره نمود که این موضوع سبب شده مقدار خطا به شدت کمتر از مقدار واقعی نشان داده شود. در خصوص ضریب تبیین (r^2) بایستی اذعان داشت که به دلیل در نظر نگرفتن مقیاس داده‌ها و دامنه تغییرات آن (از صفر تا ۱) نمی‌توان درک صحیحی از بزرگی خطا داشت. برای نمونه نمی‌توان درک صحیح و منطقی در ارزیابی خطای یک روش درون‌یابی برای داده‌های دو رقمی برای مناطق کم‌بارش کشور (مانند یزد و بم با حدود ۶۰ میلی‌متر میانگین بارش سالانه) تا چهار رقمی برای مناطق پربارش کشور (مانند انزلی و دیناچال با بارش حدود ۲۰۰۰ میلی‌متر در سال) به کمک ضریب تبیین داشت. نتایج این مطالعه نشان داد آماره MAE به دلیل استفاده از قدرمطلق و مثبت نمودن تفاوت مقادیر مشاهده شده با برآورد شده می‌تواند به عنوان معیار مناسبی برای مقایسه و انتخاب مدل دقیق‌تر مورد استفاده قرار گیرد. همچنین از نتایج این تحقیق می‌توان به عدم توافق شاخص توافق (d) با سایر آماره‌ها اشاره نمود که برای ارزیابی مدل‌های درون‌یابی پیشنهاد نمی‌شود. در نهایت آماره NSE به دلیل حساسیت به مقادیر پرت و نشان دادن مقادیر خطا بیش از مقدار واقعی پیشنهاد نمی‌شود. در یک جمع‌بندی از نتایج این تحقیق می‌توان اذعان داشت که استفاده از آماره‌های MAE و RMSE برای ارزیابی خطای روش‌های درون‌یابی به دلیل واقعی‌تر نشان دادن مقدار خطا پیشنهاد می‌شوند.

[Downloaded from system.khu.ac.ir on 2026-11-03]
[DOR: 20.1001.1.24237892.1403.11.2.7.9]
[DOI: 10.61186/jsaeh.11.2.117]

منابع

اصغری‌مقدم، اصغر؛ وحید نورانی و عطاالله ندیری. ۱۳۸۸. پیش بینی زمانی و مکانی سطح آب‌های زیرزمینی در محدوده متروی شهر تبریز با استفاده از مدل کریجینگ عصبی، تحقیقات منابع آب ایران، **۱۳:** ۲۴-۱۴.

حسنی پاک، علی اصغر؛ محمد شرف الدین. ۱۳۹۰. تحلیل داده‌های اکتشافی، مؤسسه انتشارات دانشگاه تهران.

حسینی، سیده فاطمه؛ همتی، محمد؛ جعفری، مهتاب؛ استعلاجی، علیرضا. ۱۴۰۲. تحلیل و پهنه‌بندی خطر سیل‌خیزی و ارتباط آن با پوشش گیاهی در شهرستان قیروکارزین، تحلیل فضایی مخاطرات محیطی، **۱۰(۲)**: ۷۷-۹۶.

خسروی، علیرضا؛ اژدری مقدم، مهدی؛ هاشمی‌فرد، سید آرمان؛ نظری پور، حمید. ۱۴۰۱. مقایسه نتایج تصمیم‌گیری چند معیاره در پهنه‌بندی مناطق مستعد خطر سیلاب با شاخصهای سنجش از دور در حوضه آبریز رودخانه کهیر (بلوچستان جنوبی)، تحلیل فضایی مخاطرات محیطی، **۹(۴)**: ۲۱-۴۰.

زندکریمی، آرش؛ داود مختاری. ۱۳۹۷. ارزیابی دقت روش‌های مختلف درون‌یابی در تخمین مقادیر بارش جهت انتخاب بهینه‌ترین الگوریتم در استان کردستان، پژوهش‌های جغرافیایی طبیعی، **۵(۲)**: ۳۳۸-۳۲۳.

سازمان هواشناسی کشور. ۱۴۰۰. سالنامه هواشناسی، آمار متغیرهای اقلیمی.

شمسی‌پور، علی اکبر. ۱۳۹۳. مدلسازی آب وهوایی، انتشارات دانشگاه تهران.

عیوضی، معصومه؛ مشاعدی، ابوالفضل. ۱۳۹۰پایش و تحلیل مکانی خشکسالی هواشناسی در سطح استان گلستان با استفاده از روش‌های زمین‌آماری، مرتع و آبخیزداری، **۶۴(۱)**:۶۵-۷۸.

مرادی، اسحاق؛ افسانه شهبازی؛ کاظم نصرتی؛ غلامرضا زهتابیان. ۱۳۸۵. ارزیابی روش شبیه‌سازی تصادفی برای تولید داده‌های هواشناسی، پژوهش‌های جغرافیایی، **۶۲**: ۱-۹.

ندیری، عطاالله؛ صدیقه شکور. ۱۳۹۳. ارزیابی انواع روش‌های درون‌یابی، جهت تخمین آلودگی نیترات در منابع آب زیرزمینی ، مجله هیدروژئومورفولوژی، **۱**: ۹۲-۷۵.

Alimissis, A.; k. Philippopoulos, C. Gtzanis, and D. Deligiorgi. 2018. Spatial estimation of urban air pollution with the use of artificial neural network models. *Atmospheric Environment*. **191**: 205-213.

Belkhiri L.; A. Tiri, and L. Mouni. 2020. Spatial distribution of the groundwater quality using kriging and Co-kriging interpolations. *Groundwater for Sustainable Development*. **11**: 100-73.

Chai, T.; And R. Draxler. 2004. root mean square (RMSE) or mean absolute error (MAE) Arguments against avoiding RMSE in the literature. *Geosci model Dev*. **7**: 1247-1250.

Davies, j. A.; M. Abdel-Wahab, and D. Makay.1984. Errors in estimating solar irradiance from a numerical model. *solar Energy*. **32**: 307-309.

Ding, Q.; Y. Wang, and D. Zhuang. 2018. Comparison of the common spatial interpolation methods used to analyze potentially toxic elements surrounding mining regions. *Journal of Environmental Management*. **212**: 23-31.

García-Santos, G.; M. Scheiber, and J. Pilz. 2020. Spatial interpolation methods to predict airborne pesticide drift deposits on soils using knapsack sprayers. *Chemosphere*. **258**: 127-231.

Falivene, O.; R. Cabrera, R. Tolosana-Delgado, and A. Saez. 2010. Interpolation algorithm ranking Using cross-validation and the role of asmoothing effect:A coal zone example Comput.Geosci, **36**: 512-519.

[Downloaded from system.khu.ac.ir on 2026-11-03]
[DOR: 20.1001.1.24237892.1403.11.2.7.9]
[DOI: 10.61186/jsaeh.11.2.117]

Fekete, B.; C. Vörösmarty, J. Roads, and C. Willmott .2004. Uncertainties in precipitation and their impacts on runoff estimates. *Journal of Clim*. **17**: 294–304

Fan, J.; and I. Gibels. 1996. Local Polynomial Modelling and Its Applications. *Water Resources Bulletin*. **87**: 998-1004.

Hyndman, R.; and A. Koehler. 2005. Another look at measures of forecast accuracy. *International Journal of Forecasting.* **22**: 679-688

Li, J.; H. Wan, and S. Shang. 2020. Comparison of interpolation methods for mapping layered soil particle-size fractions and texture in an arid oasis. *Catena*. **190**: 104-1014.

Jacovides, C.; G. Papaioannou, and P. Kerkides. 1994. Micro and large-scale parameters evaluation of evaporation from a lake. *Agricultural Water Management*. **13**: 263-27

Kazemi, S.; and S. Hosseini. 2011. Comparison of spatial interpolation methods for estimating heavy metals in sediments of Caspian Sea. *Expert Systems with Applications*. **38**: 1632-1649.

Moriasi, D. N.; J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel and T. L. Veith. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, **50**: 885-900.

Nekoamal, M.; and R. Mirabbasi. 2017. Assessment of interpolation methods in estimation of groundwater level (case study: sarkhon plain). *Journal of hydrogeology, online publish*. **2**: 84-95.

Nash, J.; and E. Sutcliffe. 1970. River Flow forecasting through conceptual models, part 1-A discussion of principles. *Journal of Hydrology*, **10**: 282-290.

Salah, H. 2009. Geostatistical analysis of groundwater levels in the south Al Jabal Al Akhdar area using GIS. *GIS Ostrava*, **25**: 1-10.

Willmott, C. J.; and K. Matsuura. 2005. advantages of the mean Absolute Error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, **30**: 79-82.

Willmott, C.J. 1982. Some comments on the evaluation of the model performance. *Bulletin of American Meteorological Society*, **63**: 1309-1313.

Xin, Y.; and G, Xiao 2009. Linear regression analysis: theory and computing .*world Scientific Publishin* .**348**.

Yuval-Levy, I.; and D, Broday. 2017. Improving modeled air pollution concentration maps by residual interpolation. *Science of The Total Environment*. **598**: 780-788.

[Downloaded from system.khu.ac.ir on 2026-11-03]

[DOR: 20.1001.1.24237892.1403.11.2.7.9]

[DOI: 10.61186/jsaeh.11.2.117]