

## بررسی امکان‌سنجی استفاده از پایگاه داده AgMERRA برای ساخت داده‌های ناقص و گمشده

### موجود در داده‌های ایستگاه‌های سینوپتیک

#### (مطالعه موردی: دشت مشهد)

اعظم لشکری<sup>۱</sup> - محمد بنایان<sup>۲\*</sup> - علیرضا کوچکی<sup>۳</sup> - امین علیزاده<sup>۴</sup> - یانگ سای جوی<sup>۵</sup> - ستون کی پارک<sup>۶</sup>

تاریخ دریافت: ۱۳۹۳/۰۵/۲۵

تاریخ پذیرش: ۱۳۹۴/۰۴/۱۴

#### چکیده

ساخت داده‌های ناقص (گمشده) موجود در مشاهدات ایستگاه‌های سینوپتیک با استفاده از روش‌های پیشرفته، استاندارد و دقیق در سطح جهانی اهمیت بسزایی در مطالعات هواشناسی کشاورزی و شبیه‌سازی عملکرد گیاهان زراعی دارد. لذا این پژوهش با هدف امکان‌سنجی استفاده از داده‌های AgMERRA برای تخمین داده‌های گمشده انجام شد. در این تحقیق با استفاده از خصوصیات جغرافیایی ایستگاه‌های مرجع گلمکان و مشهد داده‌های متناظر با دشت مشهد استخراج شد. با استفاده از روش‌های ارزیابی کارایی مدل‌ها شامل ضریب همبستگی، جذر میانگین مربع خطا، میانگین خطای مطلق و انحراف نتایج، واسنجی داده‌های AgMERRA انجام گرفت. نتایج واسنجی حاکی از قدرت بالای این مجموعه در شبیه‌سازی حداکثر و حداقل دمای روزانه بود. اگرچه مقادیر پارامترهای ارزیابی کارایی مدل برای بارندگی نیز در سطح قابل قبولی بود با این وجود ضریب همبستگی بارندگی روزانه بین داده‌های AgMERRA و داده‌های ایستگاهی برای ایستگاه‌های گلمکان و مشهد به ترتیب ۰/۲۵ و ۰/۴۳ بود. مقایسه توزیع احتمال تجمعی داده‌های مشاهده شده با داده‌های AgMERRA طی کل دوره مورد مطالعه برای ایستگاه‌های گلمکان و مشهد، نشان داد که داده‌های مورد استفاده از هر دو منبع از روند مشابهی تبعیت می‌کنند. با این وجود، تفاوت میانگین حداکثر دمای روزانه، حداقل دمای روزانه و بارندگی روزانه مشاهده شده و شبیه‌سازی شده طی دوره ۳۰ ساله مورد مطالعه برای ایستگاه مشهد به ترتیب ۳/۴۲ درجه سانتیگراد، ۴/۶۸ درجه سانتیگراد و ۰/۰۶ میلی‌متر در روز بود و این مقادیر برای ایستگاه گلمکان طی دوره ۲۳ ساله مورد مطالعه به ترتیب ۲/۱۰ و ۳/۰۵ درجه سانتیگراد و ۰/۲۸ میلی‌متر در روز بود. بطور کلی نتایج حاصل از این پژوهش نشان داد که برای اطمینان از دقت این داده‌ها تعداد ایستگاه‌های بیشتری مورد مطالعه قرار گیرد تا از این طریق میزان دقیق‌تری از قدرت این داده‌ها در شبیه‌سازی بارندگی نیز بدست آید.

واژه‌های کلیدی: AgMERRA، بارندگی و دما، تخمین داده‌های گمشده

#### مقدمه

(۱۴ و ۱۵). خصوصاً، اخیراً پیشرفت‌ها در زمینه مدل‌های رشد گیاهی و برنامه‌ریزی جهت کشاورزی پایدار، نیاز به داشتن پایگاه داده‌ای جهانی را به طور فزاینده‌ای افزایش داده است (۱۴). امروزه، در حقیقت، تأمین داده‌های بلند مدت آب و هوایی (داده‌های تاریخی و آینده) در مقیاس‌های مکانی و زمانی مختلف با قدرت تفکیک بالا از مهم‌ترین اولویت‌ها برای برنامه‌ریزی جهت ترویج کشاورزی پایدار است و باید در بین اولویت‌های پژوهشی کسانی قرار بگیرد که در زمینه امنیت غذایی فعالیت می‌کنند (۲).

تشعشع، حداقل و حداکثر دما و بارندگی روزانه از مهم‌ترین متغیرهای آب و هوایی مورد استفاده در مدل‌های کشاورزی و زیست محیطی هستند که توسط ایستگاه‌های هواشناسی ثبت می‌شوند. کمبود منابع مالی و سختی دسترسی به برخی مناطق، سبب شده که

مدل‌سازی‌های زیست محیطی و کشاورزی هر دو نیازمند سری کامل داده‌های هواشناسی به عنوان ورودی مدل‌ها هستند (۱۸). ثبات و شفافیت در اطلاعات آب و هوایی و روش‌های استخراج آنها، مقایسه مناطق مختلف و همچنین مقایسه مدل‌های مختلف را بخصوص زمانی که ارزیابی اثرات مورد تأکید است، تسهیل می‌کند،

۱، ۲ و ۳ - به ترتیب دانش‌آموخته دکترای بوم‌شناسی، دانشیار و استاد گروه زراعت، دانشکده کشاورزی، دانشگاه فردوسی مشهد

\* - نویسنده مسئول: (Email: banayan@ferdwosi.um.ac.ir)

۴ - استاد گروه مهندسی آب دانشکده کشاورزی، دانشگاه فردوسی مشهد

۵ و ۶ - استادیار و استاد دپارتمان مهندسی محیط زیست دانشگاه ایوا - کره جنوبی

سینوپتیک ۲۳۲۴ منطقه در نواحی مهم کشاورزی جهان می باشد. دوره آماری ۲۰۱۰-۱۹۸۰ اهمیت خاصی برای مدل های کشاورزی دارد. از این دوره برای واسنجی مدل های زراعی و اقلیمی و همچنین بعنوان دوره پایه برای تولید سناریوهای اقلیمی استفاده می شود. این مجموعه داده مانده های آب و هوایی سایر مدل های هیدرولوژیکی است با این تفاوت که در آن توجه ویژه ای به مناطق کشاورزی و متغیرهای اقلیمی مورد نیاز در مدل های رشد و نمو گیاهان زراعی شده است (۱۵).

تاکنون مطالعات متعددی در زمینه بررسی و مقایسه داده های بارندگی داده های ماهواره ای در ایران انجام شده است که می توان به مطالعات جامعی و همکاران (۱۰) در زمینه بررسی داده های بارندگی افرویدیت، غضنفری و همکاران (۷) PERSIANN، غفوریان و همکاران (۶) TRMM اشاره کرد. با توجه به اینکه پایگاه داده AgMERRA در سال ۲۰۱۴ راه اندازی شده است، مطالعات انجام شده در زمینه امکان سنجی استفاده از این مجموعه داده در سطح جهان در حال حاضر در حال انجام است و نتایج آن انتشار نیافته است و در ایران نیز هیچ مطالعه ای صورت نگرفته است.

بنابراین هدف از پژوهش حاضر ارزیابی دقت و امکان سنجی استفاده از پایگاه داده AgMERRA برای بازسازی داده های ناقص (گمشده) ایستگاه های سینوپتیک می باشد که در این مطالعه ایستگاه های گلکان و مشهد به عنوان ایستگاه های متناظر با دشت مشهد که یکی از دشت های مهم خراسان رضوی می باشد، مورد مطالعه قرار گرفتند.

## داده ها و روش مطالعه

### منطقه مورد مطالعه

منطقه مورد مطالعه در این تحقیق دشت مشهد با وسعتی بالغ بر ۱۰۰۰۰ کیلومتر مربع در شمال شرقی استان خراسان رضوی است که در حد فاصل دو رشته کوه هزار مسجد در شمال شرق و بینالود در غرب قرار گرفته است (شکل ۱). دشت مشهد از نظر شرایط آب و هوایی جزء مناطق با اقلیم نیمه خشک محسوب می شود.

در این پژوهش از داده های روزانه بارش، دمای حداکثر و حداقل ایستگاه های گلکان (دوره آماری ۲۰۱۰-۱۹۸۷ میلادی) و مشهد (دوره آماری ۲۰۱۰-۱۹۸۰ میلادی) به عنوان ایستگاه های سینوپتیک معرف دشت مشهد و به عنوان داده های مرجع استفاده شد.

### استخراج داده های AgMERRA

پس از بررسی کیفیت داده های ایستگاه های گلکان و مشهد، داده های روزانه حداکثر و حداقل دما و بارندگی Ag-MERRA برای دوره مطالعه با فرمت nc4 از پایگاه داده های سازمان فضایی آمریکا گردآوری شد (با قدرت تفکیک زمانی روزانه و مکانی  $0.25^{\circ} \times 0.25^{\circ}$ ).

در برخی مناطق دنیا ایستگاه های موجود با تراکم کمی ساخته شده باشند (۷) و بنابراین داده های ایستگاهی مکررا دارای نقصان هستند که در نتیجه آن بازسازی داده های ناقص عموماً با استفاده از روش های درون ایستگاهی، بین ایستگاهی یا مدل های آب و هوایی (آماری و فرآیند محور) و اخیراً با استفاده از فن آوری سنجش از دور صورت می گیرد (۱). علاوه بر آن، مطالعات متعدد حاکی از حساسیت بالای رشد و عملکرد شبیه سازی شده گیاهان زراعی به پارامترهای اقلیمی فصل رشد می باشد (۹). از طرفی پایگاه داده های موجود مانند افرویدیت (APHRODITE) در آسیا (۲۱)، E-Obs (۸) در اروپا تنها بارندگی را تحت پوشش قرار می دهند و در مقیاس منطقه ای هستند. همچنین برخی از پایگاه داده ها بارندگی را در مقیاس ماهانه شبیه سازی می کنند که جهت استفاده در ورودی مدل های رشد و نمو گیاهان زراعی نیازمند روش های ریز مقیاس نمایی می باشد تا تبدیل به داده های روزانه شده که انجام این تکنیک ها سبب بروز خطا و کاهش دقت نتایج می شود (۱۵). دلایل مذکور باعث شد تا پژوهش های متعددی جهت افزایش دقت و تحلیل مجدد داده های آب و هوایی روزانه مورد استفاده در مدل های رشد گیاهان زراعی آغاز شود. که از این جمله مجموعه داده ها می توان به پایگاه داده اقلیمی پریستون (۱۷)، مجموعه داده WFD (WFD-EI) برای مطالعات هیدرولوژی در سطح اروپا (۲۰) و اخیراً پایگاه داده GRASP<sup>3</sup> (۹) که توسط کشور ژاپن تهیه شده و در مطالعات امنیت غذایی کاربرد دارد، اشاره کرد.

در همین راستا، سازمان فضایی ملی ایالات متحده آمریکا<sup>۴</sup> و دانشگاه کلمبیا نیز در سال ۲۰۱۴ مجموعه داده ای را با هدف ایجاد شبکه ای جهانی و همگون برای استفاده در پژوهش های کشاورزی، امنیت غذایی و مدل های رشد گیاهی توسعه داده است. این پایگاه داده که AgMERRA نامیده شده است، مجموعه ای از داده های آب و هوایی روزانه (دمای حداکثر و حداقل، بارندگی، تشعشع، سرعت باد و فشار بخار) با قدرت تفکیک بالا برای دوره آماری ۲۰۱۰-۱۹۸۰ میلادی می باشد که به عنوان بخشی از پروژه AgMIP است و برای بررسی تأثیر تغییرات و نوسانات اقلیمی بر بخش کشاورزی تأمین شده است. این مجموعه داده حاصل تحلیل مجدد داده های ماهواره های PERSIANN، MERRA و CMORPH (۱۱) و داده های مشاهده شده از ایستگاه های

- 1 - Process-based
- 2 - Reanalysis
- 3- Global Risk Assessment toward Stable Production of Food
- 4- National Aeronautics and Space Administration (NASA)
- 5- The Agricultural Model Inter-comparison and Improvement Project
- 6 - the Modern-Era Retrospective Analysis for Research and Applications

همچنین، جذر میانگین مربع خطا (RMSE<sup>۱</sup>) (معادله ۱)، میانگین خطای مطلق (MAE<sup>۲</sup>) (معادله ۲) و انحراف نتایج (MBE<sup>۳</sup>) (معادله ۳) محاسبه شد.

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

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}| \quad (2)$$

$$MBE = \frac{(\sum_{i=1}^n (y_i - \hat{y}))}{n} \quad (3)$$

که در آن ها: n تعداد نقاط مشاهده‌ای،  $y_i$  مقدار شبیه‌سازی شده برای نقطه  $i$ امو  $\hat{y}$  مقدار مشاهده شده برای نقطه  $i$ اماست (۱۱).

چنانچه مقادیر جذر میانگین مربع خطا (RMSE) کمتر از ۱۰ باشد شبیه‌سازی عالی، بین ۱۰ تا ۲۰ خوب، بین ۲۰ تا ۳۰ متوسط و بیشتر از ۳۰ ضعیف خواهد بود (۱۹). هر چه میانگین خطای مطلق (MAE) و انحراف نتایج (MBE) نیز به سمت صفر میل کند نشان دهنده این است که مقادیر شبیه‌سازی شده تفاوت کمتری با مقادیر مشاهده شده دارند و در حقیقت مدل از کارایی بالاتری برخوردار است (۱۹). زمانی که MBE برابر صفر است، مشخص می‌شود که مدل فضای مطالعه شده را خوب برآورد کرده و هیچ‌گونه انحرافی وجود ندارد (۱۹).

علاوه بر آن برای مقایسه دقیق‌تر داده‌های مشاهده شده حداکثر و حداقل دما و بارندگی روزانه با داده‌های AgMERRA توزیع احتمال تجمعی حداکثر و حداقل دما و بارندگی روزانه نیز برای دوره آماری ۳۰ ساله در ایستگاه مشهد و دوره آماری ۲۳ ساله در ایستگاه گل‌مکان مورد بررسی قرار گرفت.

## نتایج و بحث

نتایج اعتبار‌سنجی داده‌های AgMERRA با داده‌های ایستگاه‌های سینوپتیک متناظر نشان داد که این مجموعه داده در شبیه‌سازی حداکثر و حداقل دمای روزانه و بارندگی هر دو ایستگاه مشهد و گل‌مکان از قدرت بالایی برخوردار است (جدول ۲). مقادیر جذر میانگین مربع خطا برای ایستگاه گل‌مکان برای هر سه پارامتر حداکثر و حداقل دمای روزانه و بارندگی نسبت به ایستگاه مشهد پایین‌تر بود.

در رابطه با بارندگی، اگرچه مقادیر جذر میانگین مربع خطا، میانگین خطای مطلق و انحراف نتایج کارایی بالای مدل را نشان می‌دهند اما مقادیر  $R^2$  (برای مشهد و گل‌مکان به ترتیب ۰/۴۳ و ۰/۲۵) گویای این حالت نمی‌باشد. مقدار  $R^2$  نشان می‌دهد که همبستگی بالایی بین مقادیر شبیه‌سازی شده بارندگی روزانه و مقادیر مشاهده

برای بارندگی و  $0/5^\circ \times 0/5^\circ$  برای دمای حداقل و حداکثر). داده‌ها با فرمت ذکر شده قابل ورود به محیط R می‌باشند. جدول ۱ نشان داده مشخصات ایستگاه‌های سینوپتیک گل‌مکان و مشهد می‌باشد.



شکل ۱- موقعیت مکانی منطقه مورد مطالعه  
Figure 1- Geographical characteristics of study location

جدول ۱- مشخصات ایستگاه‌های سینوپتیک

Table 1- Some characteristics of in-situ stations

	ارتفاع (m) Elevation	طول جغرافیایی (°) Longitude	عرض جغرافیایی (°) Latitude
گل‌مکان Golmakan	1176	59.17	36.29
مشهد Mashhad	999	56.36	36.16

## تعیین اعتبار داده‌های AgMERRA

معیارهای مختلفی برای ارزیابی کارایی مدل‌های مختلف وجود دارد. در این تحقیق اعتبار‌سنجی داده‌های AgMERRA برای ایستگاه‌های مذکور با استفاده از محاسبه ضریب همبستگی بین داده‌های مشاهده شده و داده‌های شبیه‌سازی شده انجام شد. ضریب همبستگی نشان‌دهنده میزان ارتباط مقادیر برآورد شده با مقادیر محاسبه شده است که هر چه مقدار آن بیشتر باشد، این ارتباط نزدیکتر است و اختلاف مقدار برآورد شده با مقدار محاسبه شده کمتر خواهد بود. ضمن اینکه برای ارزیابی بهتر، نتایج مدلها نسبت به خط ۱:۱ نیز مورد سنجش قرار گرفت.

1- Root Mean Square Error  
2- Mean Absolute Error  
3- Mean Bias Error

نتایج مقایسه میانگین داده‌های AgMERRA و داده‌های مشاهده شده در ایستگاه گل‌مکان نتایج متفاوتی را نشان داد. بطوریکه داده‌های شبیه‌سازی شده در ایستگاه گل‌مکان برای دمای حداکثر و حداقل روزانه و بارندگی به ترتیب ۲/۱۰ و ۳/۵۵ درجه سانتیگراد کمتر از داده‌های مشاهده شده برآورد شده بود و بارندگی روزانه نیز ۰/۲۸ میلی‌متر در روز بیشتر برآورد شده بود (جدول ۳). واریانس داده‌های حداکثر و حداقل دمای روزانه برای هر دو ایستگاه گل‌مکان و مشهد بسیار نزدیک به یکدیگر بود. در مورد بارندگی روزانه، واریانس داده‌های شبیه‌سازی شده در ایستگاه گل‌مکان ۰/۹۱ بیشتر از واریانس داده‌های مشاهده شده بود و برای ایستگاه مشهد ۰/۴۵ کمتر از واریانس داده‌های مشاهده شده بود، اما روند توزیع احتمال تجمعی بارندگی در هر دو ایستگاه تفاوت بسیار ناچیزی را نشان می‌دهد.

این نتایج نشان می‌دهد اگرچه کارایی مدل در شبیه‌سازی دما از قوت بالایی برخوردار بوده است، اما بررسی و مقایسه میانگین دماهای شبیه‌سازی شده و مشاهده شده طی ۳۰ سال دوره مورد مطالعه در ایستگاه مشهد و ۲۳ سال دوره مورد مطالعه در ایستگاه گل‌مکان تفاوت زیادی را بین دمای روزانه نشان می‌دهد. به عنوان مثال وجود ۳/۴۲ درجه سانتیگراد تفاوت بین حداکثر دمای روزانه مشاهده شده و شبیه‌سازی شده می‌تواند نمود مدل‌های گیاهان زراعی را در شبیه‌سازی عملکرد با استفاده از این مجموعه داده تحت تأثیر قرار دهد. چرا که در شبیه‌سازی عملکرد گیاهان زراعی نوسانات حداکثر دمای روزانه تأثیر زیادی بر عملکرد شبیه‌سازی شده خواهد داشت (۹). بطوریکه اگر تفاوت ۳/۴۲ درجه‌ای حداکثر دمای روزانه مصادف با مراحل حساس رشدی گیاهان زراعی از جمله گلدهی و پر شدن دانه باشد، نتایج کاملاً متفاوتی را به دنبال خواهد داشت.

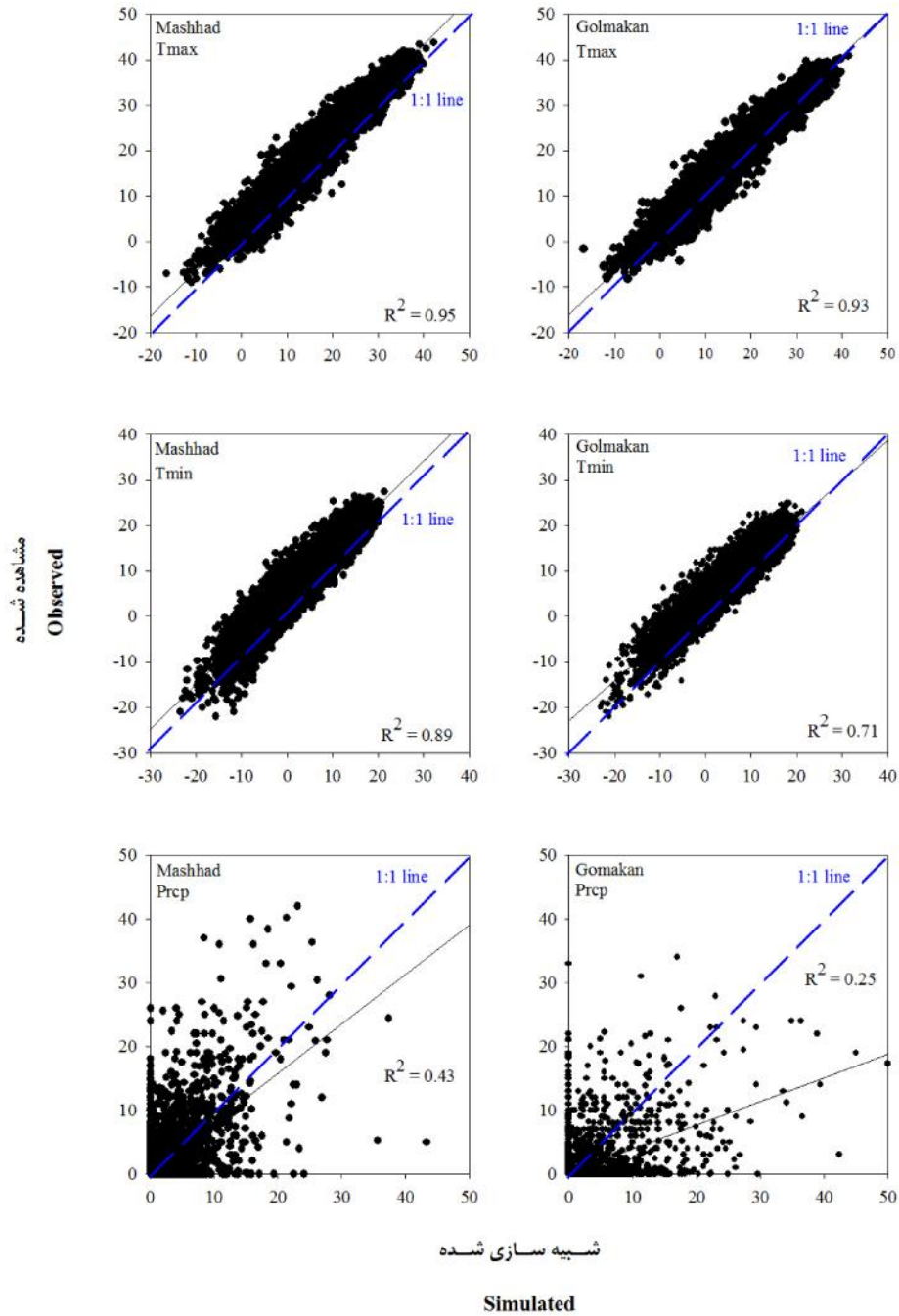
شده متناظر وجود ندارد، همچنان که خط ۱:۱ رسم شده در شکل ۲ نیز این مطلب را تأیید می‌کند. علت این پدیده را می‌توان به رفتار بارندگی ارتباط داد. چرا که خصوصیات بارندگی از جمله شدت و مقدار آن دارای تغییرات زیاد حتی در حوضه‌های کوچک می‌باشد (۱۲). بررسی رفتار بارندگی در مناطق خشک و نیمه خشک حاکی از تغییرات مکانی و زمانی زیاد بارندگی در این مناطق می‌باشد (۱۳). غضنفری و همکاران (۷) عنوان کردند بارندگی در این مناطق در بسیاری از موارد ناشی از سیستم‌های میان مقیاس و خرد مقیاس می‌باشد. در این گونه موارد بارندگی‌ها در مناطق کوهستانی از نوع اوروگرافیک بوده و در برخی موارد از نوع فرارفتی می‌باشد. بنابراین استفاده از مدل‌هایی که با استفاده از تصاویر ماهواره‌ای عمل می‌کنند می‌تواند در پیش بینی و تخمین بارندگی راهگشا باشد. اگرچه داده‌های AgMERRA تلفیقی از داده‌های ماهواره‌ای MERRA، CMORPH و PERSIANN و داده‌های جمع آوری شده از سطح ایستگاه‌های زمین مرجع می‌باشد، با این وجود مسئله عدم تخمین دقیق توزیع بارندگی هنوز به قوت خود باقی مانده است.

مقایسه توزیع احتمال تجمعی داده‌های مشاهده شده با داده‌های AgMERRA طی کل دوره مورد مطالعه برای ایستگاه‌های گل‌مکان و مشهد (شکل ۳)، نشان داد که داده‌های مورد استفاده از هر دو منبع از روند مشابهی تبعیت می‌کنند. با این وجود، میانگین حداکثر دمای روزانه در داده‌های AgMERRA برای ایستگاه مشهد به میزان ۳/۴۲ درجه سانتیگراد افزایش یافته نسبت به داده‌های مشاهده شده. همچنین میانگین حداقل دما و بارندگی روزانه نیز در داده‌های AgMERRA به ترتیب ۴/۶۸ درجه سانتیگراد و ۰/۰۶ میلی‌متر در روز کاهش یافته است (جدول ۳).

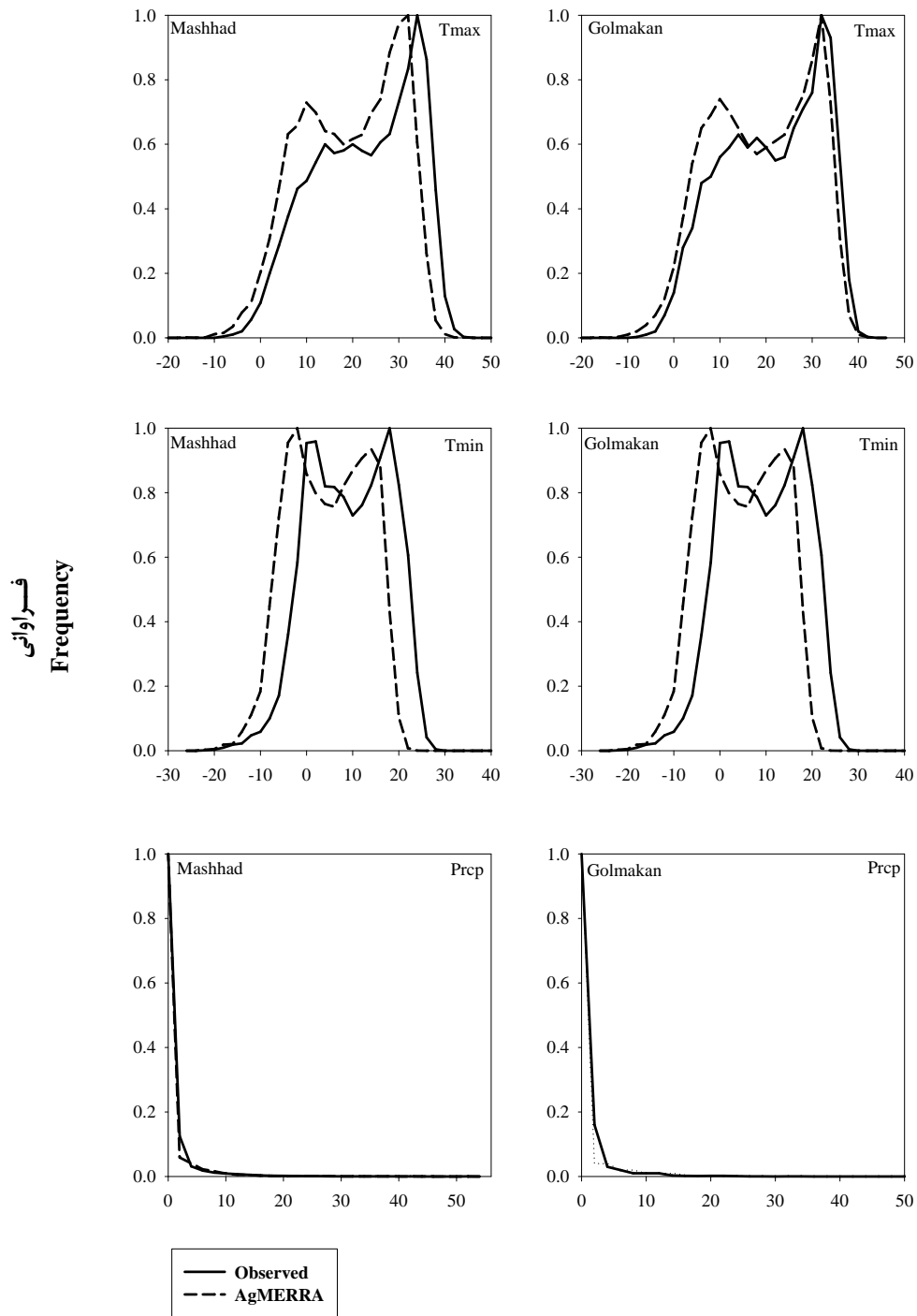
جدول ۲- مقادیر جذر میانگین مربع خطا و میانگین خطای مطلق و انحراف نتایج برای حداکثر و حداقل دما و بارندگی روزانه برای ایستگاه‌های مورد مطالعه

Table 2- RMSE, MAE and MBE for daily maximum and minimum temperature and daily precipitation in study location

ایستگاه Station	حداکثر دما (°C)			حداقل دما (°C)			بارندگی (mm)		
	Maximum Temperature			Minimum Temperature			Precipitation		
	RMSE	MAE	MBE	RMSE	MAE	MBE	RMSE	MAE	MBE
گل‌مکان Golmakan	2.95	1.65	2.06	2.85	2.06	2.97	2.56	2.67	-0.29
مشهد Mashhad	4.18	3.64	3.42	7.02	4.95	4.67	5.64	0.9	0.05



شکل ۲- مقایسه مقادیر شبیه سازی شده حداقل و حداکثر دما و بارندگی روزانه با مقادیر متناظر ایستگاه‌های گل‌مکان و مشهد  
 Figure 2- Comparison of AgMERRA data and corresponding observed data for daily maximum and minimum temperature and precipitation for Mashhad and Golmakan stations



شکل ۳- مقایسه توزیع احتمال تجمعی حداکثر و حداقل دما و بارندگی روزانه مشاهده شده با داده های AgMERRA  
 Figure 3- Comparison of probability distribution function of observed daily maximum and minimum temperature and precipitation with AgMERRA corresponding data for Golmakan and Mashhad stations

جدول ۳- میانگین و واریانس حداکثر و حداقل دما و بارندگی روزانه داده‌های مشاهده شده متناظر با داده‌های AgMERRA

Table 3- Mean and variance of observed daily maximum and minimum temperature and precipitation data and AgMERRA corresponding data

		مشاهده شده		اگمرا	
		Observed		AgMERRA	
		میانگین	واریانس	میانگین	واریانس
		Mean	Variance	Mean	Variance
حداکثر دمای روزانه (°C) Daily maximum temperature	گلمکان Golmakan	20.10	10.54	18.0	10.80
	مشهد Mashhad	21.74	10.74	18.32	10.49
حداقل دمای روزانه (°C) Daily minimum temperature	گلمکان Golmakan	6.55	7.86	3.50	8.37
	مشهد Mashhad	8.35	8.52	3.67	8.22
بارندگی (mm) Daily precipitation	گلمکان Golmakan	0.58	2.31	0.86	3.22
	مشهد Mashhad	0.70	2.82	0.64	2.37

### نتیجه گیری

های رشد و نمو گیاهان زراعی می‌باشد، لازم است نمود خروجی مدل‌های رشد و نمو گیاهان زراعی با استفاده از این مجموعه داده مورد ارزیابی قرار گیرد تا تأثیر توزیع بارندگی روزانه شبیه سازی شده و همچنین دمای روزانه بر عملکرد گیاهان زراعی با دقت بیشتری مورد بررسی قرار گیرد. بطور کلی، با توجه به اینکه ایران کشوری است با سطح زیر کشت ۱۸ میلیون هکتار بنابراین برای بررسی تغییرات عملکرد گیاهان زراعی طی سه دهه گذشته و پیش بینی تغییرات تولید در آینده، این مجموعه داده می‌تواند کمک شایانی در پر کردن داده‌های ایستگاهی گمشده موجود داشته باشد.

در این پژوهش با کمک شاخص‌های ارزیابی کارایی مدل‌ها، امکان‌سنجی استفاده از داده‌های AgMERRA جهت ساخت داده‌های گمشده (ناقص) موجود در ایستگاه‌های سینوپتیک مورد مطالعه قرار گرفت. نتایج مقایسه پارامترهای آماری  $R^2$ ، MAE، RMSE و MBE نشان می‌دهد که مجموعه داده AgMERRA قدرت خوبی در تخمین داده‌های گمشده حداکثر و حداقل دمای روزانه ایستگاه سینوپتیک دشت مشهد دارد. با این وجود، نتایج  $R^2$  مربوط به بارندگی چندان قابل قبول نبود. بنابراین لازم است در مطالعات بعدی تعداد ایستگاه‌های بیشتری مورد مطالعه قرار گیرد تا صحت این داده‌ها در سطح بزرگتری ارزیابی شود. همچنین به نظر می‌رسد برای اطمینان بیشتر از دقت داده‌های AgMERRA، مطالعه‌ای در جهت مقایسه این پایگاه داده با سایر پایگاه داده‌های موجود مانند افرودیت که خاص آسیا است، انجام شود. همچنین با توجه به اینکه هدف این مجموعه داده بکارگیری آن در اطلاعات آب و هوایی ورودی مدل-

### سپاسگزاری

نویسندگان این مقاله از کمک‌ها و راهنمایی‌های علمی دکتر الکساندر رائن پژوهشگر دانشگاه کلمبیا و سازمان فضایی ملی ایالات متحده آمریکا کمال تشکر را دارند.

### منابع

- Allen R.J., and De Gaetano A.T. 2001. Estimating missing daily temperature extremes using an optimized regression approach. *International Journal of Climatology*, 21: 1305–1319.
- Anonymous. 2013. Solutions for Sustainable Agriculture and Food Systems. Technical report for the post-2015 development Agenda. Prepared by the Thematic Group on Sustainable Agriculture and Food Systems of the Sustainable Development Solutions Network.
- Ashraf B., Yazdani R., Mousavi-Baygi M., and Bannayan M. 2013. Investigation of temporal and spatial climate variability and aridity of Iran. *Theoretical Applied Climatology*, 118: 35-46.
- Bannayan M., Mohammadian A., and Alizadeh A. 2010. On Climate Variability in North-East of Iran. *Journal of Water and Soil*, 24 (1): 118-131
- Dunn R.J.H., Willett K.M., Thorne P.W., Woolley E.V., Durre I., Dai A., Parker D.E., and Vose R.S. 2012. HadISD:

- a quality-controlled global synoptic reportdatabase for selected variables at long-term stations from 1973–2011. *Climate of the Past*, 8: 1649–1679.
- 6- Ghagourian H., and SanaeiNejad H. 2013. Drought monitoring using TRMM dataset in RazaviKhorasan Province. MSc thesis in Agrometeorology, Ferdowsi University.
  - 7- GhazanfariMoghaddam M.S., Alizadeh A., MousaviBaygi S.M., Faridhosseini A.R., and Bannayan M. 2011. Comparison the PERSIANN Model with the Interpolation Method to Estimate Daily Precipitation (A Case Study: North Khorasan). *Journal of Water and Soil*, 25 (1): 207-2015.
  - 8- Haylock M.R., Hofstra N., Klein Tank A.M.G., Klok E.J., Jones P.D., New M. 2008. A European daily high-resolution gridded data set of surface tem-perature and precipitation for 1950–2006. *Journal of Geophysical Research*. 113, D20119, <http://dx.doi.org/.1029/2008JD010201>.
  - 9- Iizumi T., Okada M., and Yokozawza M. 2014. A meteorological forcing data set for global crop modeling: Development, evaluation, and intercomparison. *Journal of Geophysical Research: Atmospheres*, 119 (2): 363–384
  - 10- Jamei M., MousaviBaygi S.M., and Bannayan M. 2014. Comparison of daily precipitation of APHRODITE network and results of spatial interpolation of precipitation in Khozestan Province. 2th National Conference on Agrometeorology.
  - 11- Joyce R.J., Janowiak J.E., Arkin P.A., and Xie P. 2004.” CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, 5: 487- 503
  - 12- Maggioni V., Sapiano M.R.P., Adler R.F., Tian Y., and Huffman G.J. 2014. An Error Model for Uncertainty Quantification in High-Time-Resolution Precipitation Products. *Journal of Hydrometeorology*, 15, 1274–1292.
  - 13- Pilgrim D.H., Chapman T.G. andDoran D.G. 1998. Problems of rainfall-runoff modeling in arid and semiarid regions. *Hydrological Sciences Journal*, 33 (4) 379-400.
  - 14- Rosenzweig C., Elliott J., Deryng D., Ruane A.C., Müller C., Arneth A., Boote K.J., Folberth C., Glotter M., Khabarov N., Neumann K., Pionteke F., Pugh T.A.M., Schmid E., Stehfest E., Yang H., and Jones J.W. 2013. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proceeding of the National Academy of Science of the United States of America*, 111(9): 3268–3273.
  - 15- Ruane A.C., Goldberg R., Chryssanthacopoulos J. 2015. Climate forcing datasets for agricultural modeling: Merged products for gap-filling and historical climate series estimation. *Agricultural and Forest Meteorology*, 200: 233-248.
  - 16- Shabani B., MousaviBaygi, M., JabariNoghabi M., and Ghareman B. 2013. Modeling and Prediction of Monthly Max &MinTemperatures of Mashhad Plain Using Time Series Models. *Journal of Water and Soil*, 27 (5): 896-906
  - 17- Sheffield J., Goteti G., Wood E.F., 2006. Development of a 50-year high-resolutionglobal dataset of meteorological forcings for land surface modeling. *Journal of Climatology*, 3088–3111
  - 18- Tardivo G., and Berti A. 2012. A Dynamic Method for Gap Filling in Daily Temperature Datasets. *Journal of Applied Meteorology and Climatology*. 51, 1079–1086.
  - 19- Wallach D., Makowski D., and Jones J.W. 2007. Working with dynamic crop models. Elsevier.
  - 20- Weedon, G. P., Gomes, S., Balsamo, G., Best, M. J., Bellouin, N., and Viterbo, P. 2012. README file for the “WFDEI” dataset. Version: September 18th, 2013, Availableat [www.eu-watch.org/data](http://www.eu-watch.org/data) availability
  - 20- Yatagai A., Kamiguchi K., Arakawa O., Hamada A., Yasutomi N., Kitoh A. 2012. APHRODITE: constructing a long-term daily gridded precipitation dataset forAsia based on a dense network of rain gauges. *Bulletin of the American Meteorological Society*, 93:1401–1415.





## Applicability of AgMERRA forcing dataset for gap-filling of *in-situ* meteorological observation, Case Study: Mashhad Plain

A. Lashkari<sup>1</sup> - M. Bannayan<sup>\*2</sup> - A. Koocheki<sup>3</sup> - A. Alizadeh<sup>4</sup> - Y. S. Choi<sup>5</sup> - S.-K. Park<sup>6</sup>

Received: 16-08-2014

Accepted: 05-07-2015

**Introduction:** Consistency and transparency in climate data and methods facilitate comparisons across regions or between models in each of these assessments, particularly when market linkages between regions are emphasized (14 and 15). However, the density and quality of stationary climate data varies widely through space and time, with the best coverage in developed countries and less reliable coverage in the Tropics and Southern Hemisphere (15). So, several groups have collected these data and constructed harmonized, global gridded datasets at monthly resolution. However, these require weather generators synthesize daily resolution before they may be applied to crop models and are therefore likely to miss events that are important for the calibration and validation of agricultural models. Regional gridded observational networks have also been created (e.g., E-Obs in Europe, (8); APHRODITE in Asia, (21)), however many regions and variables are not covered by any such network and inter comparing sites between regions with different methodologies introduces inconsistencies (). Recently, AgMERRA climate forcing dataset provide daily, high-resolution, continuous, meteorological series over the 1980–2010 period designed for applications examining the agricultural impacts of climate variability and climate change. These datasets combine daily resolution data from retrospective analyses (the Modern-Era Retrospective Analysis for Research and Applications, MERRA) with *in situ* and remotely sensed observational datasets for temperature, precipitation, and solar radiation, leading to substantial reductions in bias in comparison to a network of 2324 agricultural region stations from the Hadley Integrated Surface Dataset (HadISD) (5). Therefore, this research was done in order to investigate the possibility of using AgMERRA climate forcing dataset to estimate missing data in *in-situ* daily temperature and precipitation observations in Mashhad plain.

**Materials and Methods:** The study area was Mashhad plain in Khorasan Razavi province, located in the northeast of Iran. Climate data corresponding to Mashhad plain extracted by means of geographical characteristics of Mashhad (for the 1980-2010 periods) and Golmakan (1987-2010 period) stations from AgMERRA dataset. The goodness of fit of AgMERRA climate forcing dataset was done by means of Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Bias Error (MBE) tests and R<sup>2</sup>. The root mean-squared error (RMSE) is computed to measure the coincidence between measured and modelled values and Mean Bias Error (MBE) is simply to examine the overall model error. Furthermore, probability distribution function of observed daily data and AgMERRA data for both Golmakan and Mashhad stations calculated. Eventually, mean and variance of AgMERRA and *in-situ* data were calculated to have a more accurate comparison of simulated and observed data.

**Results:** Results indicated that AgMERRA dataset has a good performance in estimating daily maximum and minimum temperature in Mashhad Plain. RMSE, MAE and MBE for daily precipitation illustrated a good performance of AgMERRA data. However, R<sup>2</sup> value was 0.43 and 0.25 for Mashhad and Golmakan stations, respectively. Although the probability distribution function of daily maximum and minimum temperature and precipitation indicated the same trend for both studied stations, comparison of mean and variance of observed daily maximum and minimum temperature and precipitation and AgMERRA data for Mashhad and Golmakan stations showed different results. The difference between mean of AgMERRA and observed daily maximum temperature for Mashhad and Golmakan stations was 3.42 and 2.10 °C, respectively. It was 4.68 and 3.05 °C for minimum daily temperature for Mashhad and Golmakan, respectively, and the difference between mean of AgMERRA and observed daily precipitation was 0.06 and 0.28 mm.day<sup>-1</sup> for Mashhad and Golmakan, respectively.

**Discussion and Conclusion:** This research showed that using AgMERRA climate forcing dataset could be a

1, 2 and 3- Ph.D in Agroecology, Associated Professor and Professor of Agronomy Department, Agriculture Faculty, Ferdowsi University of Mashhad, Iran, Respectively

(\*Corresponding Author Email: banayan@ferdowsi.um.ac.ir)

4- Professor of Water Engineering Department, Agriculture Faculty, Ferdowsi University of Mashhad, Iran

5 and 6- Assistant Professor and Professor of Environmental Engineering Department, Ewha Womans University, South Korea, Respectively

reliable tool to estimate missing data of *in-situ* temperature observations. Although the performance of AgMERRA dataset was good for daily precipitation, distribution of simulated precipitation compare with observed precipitation was different. Concerning AgMERRA precipitation data some points have to keep in mind that precipitation in arid and semi-arid regions tends to be more variable in time than in humid regions. In fact, the distinctive features of arid and semiarid regions affect precipitation modeling on a discrete event basis and a continuous basis (7, 10, 13). Results of this research illustrated the same trend and it revealed that AgMERRA dataset could not simulate the precipitation distribution in Mashhad plain. It seems that comparing AgMERRA precipitation data with OPHRODITE dataset and other dataset can give us more accurate vision about AgMERRA dataset. Furthermore, it seems that it is needed to do more researches regarding investigation of performances of crop model results by using AgMERRA dataset as climate data input, because this dataset was released for agricultural application.

**Keywords:** AgMERRA, Estimating Missing Data, Precipitation- Temperature