

بازسازی محصولات مفقوده کلروفیل a مودیس با استفاده از الگوریتم DINEOF در نرم افزار R: مطالعه موردی خلیج فارس و دریای عمان

مریم کریمیان^۱، امید بیرقدار کشکولی^{۱*}، رضا مدرس^۱، سعید پورمنافی^۱

۱- دانشکده منابع طبیعی، دانشگاه صنعتی اصفهان، اصفهان، ایران

چکیده

هدف این مطالعه، استفاده از الگوریتم DINEOF برای پر کردن نقاط فاقد داده در تصاویر ماهواره‌ای غلظت کلروفیل a در آبهای سطحی خلیج فارس و دریای عمان می‌باشد. در این مطالعه از داده‌های سطح ۳ (وضوح مکانی ۴ کیلومتر) غلظت کلروفیل a حاصل از سنجنده مودیس در بازه زمانی ۲۰۰۳ تا ۲۰۲۰ در خلیج فارس و دریای عمان (استخراج شده از سایت ناسا) استفاده شد. برخی از تصاویر در بعضی از ماه‌های سال دارای نقاط فاقد محصول بودند. تصاویر دارای محصولات مفقوده (گپ یا شکاف) غلظت کلروفیل a در خلیج فارس و دریای عمان با استفاده از بسته نرم افزاری $rtsa.gapfill$ و الگوریتم DINEOF در نرم افزار R بازسازی شدند. در نهایت برای ارزیابی صحت‌سنجی الگوریتم DINEOF، پارامترهایی نظیر $RMSE$ ، MSE ، MAD و SNR محاسبه شده است. نقشه‌های اصلی (دارای شکاف) غلظت کلروفیل a نشان دادند که بیشترین میزان محصولات مفقوده در ماه جولای ۲۰۰۹ و ۲۰۱۵ وجود دارد. با ارزیابی صحت بازسازی مشخص شد بازسازی داده‌های مفقوده به خوبی توسط الگوریتم صورت گرفته است (به عنوان مثال در ماه جولای ۲۰۱۴ پارامترهای محاسبه شده بدین شرح می‌باشند: $R^2 = 0.83$ ، $RSME = 0.34$ ، $MAD = 0.14$ و $MSE = 0.10$). بررسی نتایج نشان داد که اجرای الگوریتم DINEOF در نرم افزار R برای بازسازی محصولات مفقوده موجود در تصاویر غلظت کلروفیل a روشی کارآمد، سریع و موفقیت‌آمیز است. محصولات کلروفیل a حاصل از بازسازی برای کاربردهایی مانند تشخیص شکوفایی جلبکی و همچنین مدیریت مناسب شیلاتی قابل استفاده است. در مقیاس زمانی بلندمدت، مقایسه محصولات مفقوده با بازسازی شده به کنترل کیفیت داده‌های ماهواره‌ای کمک می‌کند.

کلیدواژه‌ها: غلظت کلروفیل a ، مودیس، بازسازی محصولات، DINEOF، خلیج فارس، دریای عمان

مقدمه

فیتوپلانکتون‌ها پایه زنجیره غذایی دریایی هستند و به حفظ زندگی در اقیانوس‌ها از طریق فتوسنتز کمک می‌کنند. آن‌ها همچنین برای زندگی بر روی زمین نیز (به دلیل تولید بیش از نیمی از ذخیره اکسیژن زمین) مهم هستند. فیتوپلانکتون‌ها نقش اساسی در چرخه کربن دارند، زیرا آنها دی اکسید کربن اقیانوس را در طول فتوسنتز مصرف و اکسیژن را به عنوان یک محصول جانبی منتشر می‌کنند. نظارت بر میزان، وسعت و پراکنش فیتوپلانکتون‌ها مورد توجه دانشمندان، محققان و مدیران منابع آبی برای چندین دهه بوده است. شناخت موجودات فیتوپلانکتونی دریایی و پراکنش آنها، محققان را قادر می‌سازد تا در مورد سلامت پیکره آب، ترکیب و وضعیت اکولوژیک آن نتیجه‌گیری کنند [۱]. برای نظارت، ارزیابی و در نهایت گام نهادن در مدیریت پایدار منابع دریایی [۲]، ارائه اطلاعات طولانی‌مدت در مورد پویایی فیتوپلانکتون‌ها ضروری است. با این حال، اندازه‌گیری‌های درمکان فیتوپلانکتون‌ها به دلیل صرف زمان و هزینه بالا در مقیاس وسیع دریاها و اقیانوس‌ها امکان پذیر نمی‌باشد. از این‌روی، حسگرهای ماهواره‌ای رنگ اقیانوس به عنوان ابزاری توانمند و کارآ برای نظارت بر محیط دریایی استفاده می‌شوند [۳]. طیف‌سنجی کلروفیل دریا بر اساس تصاویر ماهواره‌ای در طیف مرئی (طول موج بین ۴۰۰ تا ۷۰۰ نانومتر)، یک روش مفید برای ارزیابی فراوانی غلظت کلروفیل a است [۴] و [۵] و برای برآورد غلظت کلروفیل a به منظور برآورد فیتوپلانکتون‌ها استفاده می‌شود [۶]. داده‌های ماهواره‌ای اغلب به دلیل عواملی مانند وجود

نوع مقاله

مقاله پژوهشی اصیل

تاریخ دریافت: ۱۴۰۰/۰۷/۲۰

تاریخ پذیرش: ۱۴۰۱/۲/۱۵

تاریخ چاپ الکترونیکی: ۱۴۰۱/۰۳/۳۰

*نویسنده مسول:

omid.beyraghdar@iut.ac.ir

ابرها و رد مشاهدات به وسیله‌ای کنترل کیفیت دارای گپ هستند. در برخی از داده‌های ماهواره‌ای خاص که توسط بعضی از رادیومترها تولید می‌شوند مشکل حضور پوشش ابر وجود ندارد، اما ممکن است پوشش آن‌ها در معرض محدودیت‌های دیگر مانند ماسک ساحلی یا آلودگی باران باشد [۷]. تکنیک‌های بازسازی داده‌های از دست رفته برای مجموعه داده‌های رنگ اقیانوس نسبت به آن‌هایی که برای دمای سطح آب استفاده می‌شوند، کمتر توسعه یافته‌اند. شاید به این دلیل که داده‌های ماهواره‌ای کلروفیل به تازگی و نسبتاً به راحتی در دسترس قرار گرفته‌اند و صحت بازیابی کلروفیل نسبت به دمای سطح آب و برخی از پارامترهای دیگر ضعیف است. در سال‌های اخیر، در چندین مطالعه سعی در بازسازی داده‌های ناقص کلروفیل با استفاده از روش‌های مختلف به عنوان مثال کریجینگ [۸] و سایر روش‌های (Optimal Interpolation) OI شده است [۹]. آنالیز OI برای محاسبه غالب‌ترین تابع متعامدتجربی استفاده می‌شود و سپس این توابع متعامدتجربی در داده‌های درمکان قرار می‌گیرند. یکی از عیوب روش درون‌یابی بهینه به عنوان روشی برای بازسازی داده‌های از دست رفته این است که اطلاعات قبلی برای آمار خطای داده‌ها ضروری است و به طور کلی این روش ضعیف شناخته شده است [۱۰]. هزینه محاسباتی این تکنیک‌ها ممکن است هنگام استفاده از ماتریس‌های بزرگ بالا باشد [۱۱]. کریجینگ روشی مبتنی بر رگرسیون است که از نقاط داده‌های اطراف مشاهده شده استفاده و با توجه به مقادیر کوواریانس وزن می‌شود. به طور کلی روش کریجینگ فقط در مورد داده‌های مکانی اعمال می‌شود که داده‌های مکانی مربوط به مکان داده‌ها و فاصله بین نقاط موقعیت در نظر گرفته شده است. در مطالعات اخیر روش نوینی را برای پر کردن داده‌های از دست رفته با استفاده از یک الگوریتم مبتنی بر توابع متعامدتجربی به نام (Data Interpolating Empirical Orthogonal Functions) DINEOF توصیف شده است [۱۲، ۱۳] که این روش برای اولین بار در مورد داده‌های دمای سطح آب دریای آدریاتیک به کار برده شده است. در این مطالعات یک روش خودسازگار جدید برای استنباط داده‌های از دست رفته از مجموعه داده‌های اقیانوس‌شناسی و برای استخراج توابع متعامدتجربی مربوطه ارائه شده است. روش جدید به عنوان یک محصول جانبی، امکان تشخیص تعداد EOF، نویز و خطای درون‌یابی را فراهم می‌کند [۱۳، ۱۴]. همچنین در مطالعه‌ای دیگر الگوریتم DINEOF برای عناصر چند متغیره برای بازسازی دمای سطح آب دریا، غلظت کلروفیل و باد در ساحل غربی فلوریدای ایالات متحده آمریکا به صورت موفقیت‌آمیز مورد استفاده قرار گرفت [۷]. الگوریتم DINEOF به وسیله‌ی گروه (GeoHydrodynamics and Environment Research) GHER توسعه یافته است. پایه‌های این روش شامل یک روش تکرار شونده (Iterative) برای محاسبه مقادیر فیلدی در مکان‌های بدون داده است. در این الگوریتم ابتدا مقدار میانگین مکانی و زمانی از آنالیز داده‌ها حذف می‌شود و داده‌های از دست رفته با عدد صفر مقداردهی اولیه می‌شوند. سپس یک تابع متعامدتجربی از این فیلد محاسبه می‌شود، داده‌های از دست رفته با مقادیر به دست آمده توسط تجزیه تابع متعامدتجربی جایگزین می‌شوند و این الگوریتم تا همگرایی مقادیر داده‌های از دست رفته تکرار می‌شود. بعد از آن، دو تابع متعامدتجربی محاسبه می‌شود و کل روش تکرار می‌شود و سپس به دنبال آن سه تابع متعامدتجربی و به همین ترتیب مراحل ادامه می‌یابد. تعداد کل تابع متعامدتجربی‌های محاسبه شده به وسیله اعتبارسنجی متقابل تعیین می‌شوند، به شکلی که چند داده معتبر (معمولاً ۱ درصد از کل داده‌ها) در ابتدا از مجموعه داده‌ها جدا شده و برای اعتبارسنجی مد نظر قرار می‌گیرد. در هر تکرار تابع متعامدتجربی، مقدار محاسبه شده از سری تابع متعامدتجربی با مقدار واقعی این داده‌های کنار گذاشته شده، مقایسه می‌شود. تعداد توابع متعامدتجربی که این اختلاف را به حداقل می‌رساند به عنوان تعداد مطلوب برای بازسازی مجموعه داده‌ها حفظ می‌شود [۱۴]. مزیت استفاده از الگوریتم DINEOF نسبت به سایر روش‌ها مانند درون‌یابی بهینه این است که الگوریتم DINEOF اطلاعات مورد نیاز برای بازسازی را از تغییرات خود داده‌های در دسترس کسب می‌کند. بنابراین هیچ عامل خارجی و قضاوت شخصی در محاسبه پارامترها دخیل نیست [۱۵، ۱۶] و همچنین از نظر زمان محاسباتی سریعتر از OI است [۱۳]. سایر روش‌های درون‌یابی سنتی مانند نزدیکترین همسایه یا درون‌یابی دو خطی (Bilinear interpolation) از اطلاعات مکانی در تصاویر به طور کامل استفاده نمی‌شود. صفحه اصلی سایت <http://modb.oce.ulg.ac.be/mediawiki/index.php/DINEOF> شامل دستورالعمل‌های اساسی برای دانلود و تدوین DINEOF است. کد منبع و فایل‌های باینری (فایل‌های غیر متنی یا ۰ و ۱) بطور رایگان در این صفحه قابل دانلود کردن هستند. به طور کلی DINEOF شامل یک سری روال‌های (Routines) فورترن ۹۰ است که مجتمع شده‌اند و تحت مجوز عمومی گنو به

طور رایگان در دسترس هستند. بسته کامل آن در <http://modb.oce.ulg.ac.be/mediawiki/index.php/DINEOF> قابل دانلود کردن می‌باشد. در این صفحه وب، آخرین نسخه نرم افزار یافت می‌شود و توسط توسعه دهندگان و از جمله جدیدترین تحولات آزمایش شده است و این روال‌ها قابل تنظیم و متناسب با نیاز کاربر است [۱۴]. نرم افزار ARPACK [۱۶] برای حل مشکلات مقدار ویژه (Eigen Value) مورد نیاز است و در صورت نیاز به پشتیبانی netCDF، کتابخانه netCDF نیز باید نصب شود. برای آن دسته از کاربران که تجربه کم و یا اصلاً تجربی در تدوین برنامه‌ها از کد منبع ندارند، یک بسته آماده برای استفاده که از قبل تدوین شده، نیز موجود است. این بسته برای سیستم عامل‌های لینوکس و ویندوز وجود دارد و نیازی به دانش فورترن نیست [۱۴]. نرم افزار DINEOF قابل نصب در سیستم عامل لینوکس و ویندوز می‌باشد ولی نصب آن به سختی انجام می‌گیرد. در حال حاضر هیچگونه پیاده‌سازی رسمی در نرم افزار R وجود ندارد اما کد R برای DINEOF به عنوان بخشی از بسته sinkr تحت توسعه موجود است. استفاده از نسخه پیچیده رسمی DINEOF از نظر محاسباتی برای مناطق مطالعه‌ای با وسعت زیاد امکان پذیر نیست، بنابراین می‌توان از نرم افزار R استفاده کرد. این نرم افزار به صورت رایگان در دسترس است و همچنین این محیط محاسباتی برای بازسازی محصولات به وسیله الگوریتم DINEOF خیلی سریع عمل می‌کند [۱۷]. در مطالعه‌ای از الگوریتم DINEOF برای تصاویر ۸ روزه غلظت کلروفیل a حاصل از داده‌های ماهواره‌ای برای تولید روندهای طولانی مدت مجموعه داده‌های کلروفیل a بدون ابر در دریای بوهایی و دریای زرد چین در بازه زمانی ۱۹۹۷ تا ۲۰۱۰ به کار برده شد [۱۸]. در مطالعه دیگر از یک سری زمانی سه ساله (۲۰۱۴ تا ۲۰۱۶) برای ارزیابی دقت خروجی بازسازی تصاویر کلروفیل a حاصل از سنجنده مودیس و ماهواره آکوا با الگوریتم DINEOF برای دریای سالیس کانادا استفاده کردند [۱۹]. محققانی از الگوریتم DINEOF برای بازسازی محصولات غلظت کلروفیل a سطح دریای چین شرقی در بازه زمانی ۲۰۰۳ تا ۲۰۱۶ و همچنین از تجزیه و تحلیل همبستگی و رگرسیون برای ارزیابی رابطه بین غلظت کلروفیل a و دمای سطح دریا استفاده کردند [۲۰]. در مطالعه‌ای با به کارگیری الگوریتم DINEOF به بازسازی تصاویر ماهانه سطح ۳ (۴ کیلومتر) کلروفیل a و دمای سطح آب حاصل از مودیس - آکوا خلیج فارس پرداختند [۲۱]. در تحقیقی محصولات از دست رفته دمای سطح آب ماهواره‌ای مودیس و سنجنده آکوا از ژانویه ۲۰۰۳ تا مارس ۲۰۱۵ در دریای چین شرقی را توسط الگوریتم DINEOF بازسازی کردند و نتایج بازسازی با تجزیه و تحلیل خطا قبل و بعد از اعتبارسنجی متقابل بازسازی محصولات ارزیابی شدند. در این مطالعه مقدار دو شاخص (Root-Mean-Squared Error) RMSE و (Mean absolute error) MAE به ترتیب ۰/۴۲۷۲ و ۰/۵۹۵۳ بدست آمده که نشان دهنده‌ی اثر بخش بودن نتایج بازسازی می‌باشد [۲۲]. پروژه‌ای تحت عنوان *Ecopotential: Improving Future Ecosystem Benefits Through Earth Observations* و همچنین با عنوان فرعی *Time series EOFs (map and analysis)* انجام گرفت که در بخش‌هایی از آن به بررسی روش‌های مختلف برای پرکردن نقاط فاقد داده پرداخته شده است [۲۳]. یکی از روش‌های ذکر شده در این پروژه الگوریتم DINEOF بوده است که تحت عنوان تابع *rtsa.gapfill* در نرم افزار R توسط یکی از محققین این پروژه (Filipponi) آماده‌سازی شده است. این بسته نرم‌افزاری در صفحه اینترنتی به آدرس (<https://rdr.io/github/ffilipponi/rtsa/man/rtsa.gapfill.html>) قابل دسترس می‌باشد [۲۳]. در مطالعه‌ای دیگر محققان به بررسی تفاوت الگوریتم‌های پر کردن ابر برای محصولات دریایی پرداختند. در این مطالعه از نرم افزار R برای مناطق با وسعت بیشتر و از نرم افزار رسمی DINEOF برای مناطق با وسعت کمتر استفاده شده است [۱۷]. نتایج مطالعات نشان می‌دهد که الگوریتم DINEOF را می‌توان با موفقیت روی متغیرهای بسیار با ویژگی‌های مختلف اعمال کرد، اما تغییرات خاص با توجه به ویژگی مجموعه داده می‌تواند دقت داده‌های بازسازی شده را بهبود بخشد [۲۴]. نتایج مطالعات اشاره شده همچنین کارایی و عملکرد مناسب الگوریتم DINEOF را برای بازسازی محصولات دمای سطح آب و غلظت کلروفیل a و غیره مورد تایید قرار داده‌اند و به عنوان شناخته‌ترین و معمول‌ترین روش بازسازی محصولات کلروفیل مورد تایید دانشمندان حوزه سنجش از دور و اقیانوس‌شناسی می‌باشد. در مطالعه حاضر با بررسی‌های انجام شده، برای اولین بار از داده‌های با وضوح بالا (۴ کیلومتر) و طولانی مدت (۲۰۰۳ تا ۲۰۲۰) و الگوریتم کد نویسی شده در محیط نرم افزار Rstudio استفاده شده است که به صورت جامع کل

منطقه مورد مطالعه، مورد بررسی و اعتبارسنجی قرار داده شده است. هدف از این مطالعه بازسازی محصولات از دست رفته کلروفیل *a* در خلیج فارس و دریای عمان در بازه زمانی ۲۰۰۳ تا ۲۰۲۰ توسط الگوریتم DINEOF در زبان برنامه‌نویسی R می‌باشد.

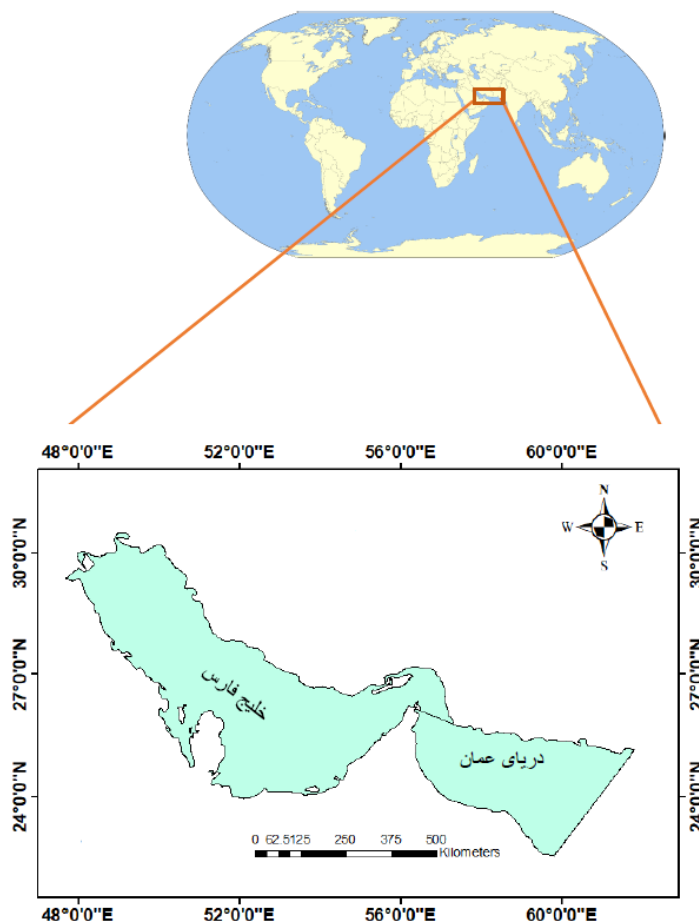
مواد و روش‌ها

منطقه مورد مطالعه

در مرزهای جنوبی ایران دو اکوسیستم آبی خلیج فارس و دریای عمان (شکل ۱) قرار دارد. خلیج فارس یک دریای حاشیه‌ای نیمه محصور است که به وسیله توده‌ی خشکی احاطه شده است. این خلیج بین ۲۴ تا ۳۰ درجه عرض شمالی و ۴۸ تا ۵۷ درجه طول شرقی قرار دارد و با حداکثر طول و عرض ۱۰۰۰ و ۳۷۰ کیلومتر است و دارای میانگین عمق ۳۵ متر در بیشتر سواحل عربی و عمق ۶۰ متر در سواحل ایران می‌باشد [۲۵]. دریای عمان بین ۲۲ درجه تا ۲۶ درجه شمالی و ۵۶ درجه تا ۶۲ درجه شرقی قرار دارد. این دریا ۵۴۵ کیلومتر طول [۲۶] و مساحتی برابر با ۹۴۰۰۰ کیلومتر مربع دارد [۲۷]. اقلیم آن تحت تأثیر دو فصل مونسونی یعنی مونسون زمستانی (از نوامبر تا آوریل رخ می‌دهد) که با بادهای کم سرعت شمال شرقی همراه است و مونسون تابستانی (از ماه ژوئیه تا سپتامبر) که با بادهای جنوبی و جنوبی غربی با سرعت بالا همراه است، قرار می‌گیرد [۲۸]. نتایج مطالعات انجام شده در خلیج فارس نشان داده است پراکنش غلظت کلروفیل *a* در نیمه شرقی خلیج فارس بیشتر از نیمه غربی آن بوده به طوریکه متوسط سالانه آن برای بخش شرقی خلیج فارس ۰/۵۶ ولی برای بخش غربی ۰/۴۲ میلی‌گرم در متر مکعب گزارش شده است [۲۹]. مطالعات انجام شده در خلیج فارس میانگین غلظت کلروفیل *a* را برای این پهنه آبی ۰/۳۸ میلی‌گرم در متر مکعب گزارش کرده‌اند [۳۰]. میانگین دامنه تغییرات کلروفیل *a* در کل آب‌های سطحی خلیج فارس بین ۰/۲ تا ۰/۸ و دریای عمان پس از مونسون بین ۲ تا ۲۰ میلی‌گرم بر متر مکعب گزارش شده است [۳۱]. همچنین نتایج مطالعات به عمل آمده در محدوده دریای عمان، نوسانات غلظت کلروفیل *a* را در لایه‌های سطحی این پهنه آبی بین ۰/۱ تا ۳ میلی‌گرم در متر مکعب گزارش نموده و نشان داده است که غلظت آن در مناطق ساحلی به مراتب بیشتر از مناطق دریایی می‌باشد [۳۰، ۳۱].

محصولات

در این مطالعه محصولات ماهانه غلظت کلروفیل *a* سطح ۳ (۴ کیلومتر) سنجنده مودیس از ژانویه ۲۰۰۳ تا دسامبر ۲۰۲۰ از سایت Ocean Color ناسا (<http://oceancolor.gsfc.nasa.gov/>) برای منطقه خلیج فارس و دریای عمان دانلود شد (شیپ فایلی با مشخصات جغرافیایی ۳۰/۷۱-۲۲/۰۳ درجه عرض شمالی و ۷۰/۹۷-۴۷/۲۹ درجه طول شرقی). در بعضی از ماه‌های سال به خصوص ماه‌های گرم سال به دلیل عواملی مانند پوشش ابر و عدم برداشت داده توسط سنجنده، انقطاع و گسست‌های داده‌ای (شکاف) در این تصاویر وجود داشت. در این مطالعه تصاویر ماه جولای برای بعضی از سال‌ها (۲۰۰۷ تا ۲۰۱۸) در محدوده زمانی مورد مطالعه به دلیل وجود نقاط زیادی که فاقد داده هستند به عنوان نمونه آورده و برای بازسازی محصولات مفقوده از الگوریتم DINEOF استفاده شده است. در این مطالعه تمام لایه‌ها (۲۱۶ لایه) و یک لایه ماسک در محیط R برای بازسازی تصاویر غلظت کلروفیل *a* آورده و بدین منظور از تابع `rtsa.gapfill` و همچنین الگوریتم DINEOF برای بازسازی سری‌های زمانی رستری استفاده شده است.



شکل ۱. منطقه مورد مطالعه

تابع `rtsa.gapfill` (Raster time series gapfilling)

تابع `rtsa.gapfill` در نرم افزار R توسط Filipponi یکی از محققین پروژه Ecopotential: Improving Future Ecosystem Benefits Through Earth Observations آماده‌سازی شده [۲۳] و به صورت متن‌باز در صفحه اینترنتی به آدرس <https://rdrr.io/github/ffilipponi/rtsa/man/rtsa.gapfill.html> قابل دسترس می‌باشد. مراحل انجام کد `rtsa.gapfill` در محیط نرم‌افزار R به شرح زیر است:

- ۱- ایجاد `rtts` (raster time series): با استفاده از دستور `stack`، داده‌های به فرمت `netCDF` به شکل سری‌های زمانی رستری در می‌آید.
- ۲- تولید ماسک رستری: یک ماسک، پیکسل‌های قابل قبول را برای بازسازی شناسایی می‌کند. لایه‌های ماسک برای تشخیص پیکسل‌های خشکی از پیکسل‌های دریا و برای جدا کردن پیکسل‌های منفرد اقیانوس موجود در کمتر از ۲٪ تصاویر ماهواره‌ای کلروفیل تعریف شده است. این ماسک‌ها برای شناسایی پیکسل‌های معتبر دریایی مشترک در همه محصولات ورودی یکپارچه استفاده می‌شوند [۱۹].
- ۳- مقادیر به طور تصادفی از سلول‌های شیء `ITS` حذف می‌شوند.
- ۴- تعیین موقعیت تصادفی سلول‌ها که با `NaN` تنظیم می‌شود.
- ۵- پر کردن نقاط فاقد داده در مجموعه داده‌های دارای انقطاع یا گسست داده‌ای.

بسته `rtsa`

بسته *rtsa* (Raster Time Series Analysis) برای زبان برنامه‌نویسی *R*، مجموعه‌ای از آنالیزها برای انجام آنالیز مکانی-زمانی از سری‌های زمانی رستری و دسترسی رایگان برای زبان برنامه‌نویسی *R* می‌باشد. بسته *rtsa* به عنوان یک پیش‌فرض برای توابع موجود در بسته‌های مختلف نرم‌افزار *R* (که به طور خاص برای کنترل مجموعه داده‌های جغرافیایی ارائه شده به عنوان سری‌های زمانی رستری طراحی شده‌اند) عمل می‌کنند. در این مطالعه همانطور که در بخش‌های قبل گفته شده است با استفاده از دستور *stack* داده‌های با فرمت *netCDF* به سری‌های زمانی رستری تبدیل شدند.

توابع موجود در بسته، اجازه ورود مستقیم سری‌های زمانی رستری برای استخراج اطلاعات مختصر و جامع با استفاده از مزایای بسیاری از تکنیک‌های آنالیزی را می‌دهد. موفقیت تیم *ISPRa* (یک موسسه تحقیقاتی دولتی است که توسط وزارت محیط‌زیست در ایتالیا تأسیس شده است که در زمینه استفاده پایدار از محیط‌دریایی، نظارت و کنترل بر آن مشاوره ارائه می‌دهد) در طی تجزیه و تحلیل سری‌های زمانی معطوف به مدیریت مجموعه داده‌های جغرافیایی ارائه شده به عنوان سری زمانی رستری برای استخراج اطلاعات جامع برای ارزیابی خدمات اکوسیستم است. ویژگی‌های اصلی این بسته عبارتند از:

- استفاده از سری‌های زمانی رستری با بُعد زمانی واضح
- استفاده از سری‌های زمانی رستری به عنوان ورودی مستقیم با توابع
- استفاده از ماسک رستری برای انتخاب منطقه مورد نظر و کاهش بار حافظه
- پردازش موازی با استفاده از *CPU* های متعدد
- در این بسته از روش‌های زیر برای بازسازی محصولات مفقوده استفاده می‌شود [۲۳].

DINEOF

درون‌یابی *line*: روشی برای برازش منحنی با استفاده از چند جمله‌ای‌های خطی برای پیش‌بینی مقادیر از دست رفته در سری‌های زمانی می‌باشد. درون‌یابی *spline*: یکی دیگر از روش‌های پر کردن انقطاع داده‌ها به صورت یک بُعدی است که در آن درون‌یاب، نوع خاصی از یک چندجمله‌ای تکه‌ای است که یک *spline* نامیده می‌شود. درون‌یابی *stine*: یک روش دیگر درون‌یابی یک بُعدی است که مبتنی بر چند جمله‌ای‌ها می‌باشد و نتیجه آن شیب کمتر در نقاط دارای تغییرات ناگهانی در سری زمانی می‌باشد که در نهایت احتمال داشتن داده‌های خارج از محدوده را کاهش می‌دهد [۲۳]. در این مطالعه از الگوریتم *DINEOF* برای بازسازی محصولات مفقوده کلروفیل *a* استفاده شده است که توضیح آن به شرح زیر است:

الگوریتم DINEOF

الگوریتم *DINEOF* یک روش خودسازگار (Self-adaptive) برای بازسازی داده‌های از دست رفته در محصولات اقیانوس‌شناسی است که به اطلاعات قبلی نیاز ندارد. این الگوریتم مبتنی بر تابع متعامدتجربی است که برای محصولات از دست رفته در متغیرهای مختلف اما مرتبط استفاده می‌شود [۲۳]. الگوریتم *DINEOF* به شرح مراحل زیر در این مطالعه استفاده شده است [۱۸].

مرحله اول: محصولات اصلی در ماتریس اولیه (*X*) با ابعاد $m \times n$ ذخیره می‌شود. *m* تعداد پیکسل‌ها و *n* تعداد تصاویر است. این ماتریس شامل هر دو داده موجود و مفقوده است. داده‌های مفقوده به شکل *NaN* نمایش داده می‌شود. ابتدا ۵ درصد از داده‌های موجود در ماتریس به صورت تصادفی انتخاب شده و برای اعتبارسنجی پیشرفته کنار گذاشته می‌شوند. سپس مقدار میانگین زمانی مکانی از ماتریس کم می‌شود و داده‌های از دست رفته (از جمله اعتبارسنجی متقابل داده) برای تضمین اینکه آن‌ها نسبت به *X* بی‌طرف هستند با صفر مقادردهی می‌شوند.

مرحله دوم: برای بازسازی داده‌های از دست رفته، تجزیه تابع متعامدتجربی توسط روش *SVD* (Singular Value Decomposition) محاسبه می‌شود و توابع متعامدتجربی مکانی (*U*)، ماتریس مقادیر واحد (*S*) و توابع متعامدتجربی زمانی (*V*) بدست می‌آید. داده‌های مفقوده را می‌توان توسط توابع متعامدتجربی کوتاه شده (Truncated EOFs) بازسازی کرد:

$$X_{i,j} = \sum_{p=1}^K S_p (U_p) (V_p^T)$$

در این فرمول $X_{i,j}$ داده‌های از دست رفته، U_p و V_p به ترتیب ستون p ام توابع متعامدتجربی مکانی و زمانی، S_p مقدار ویژه p ام و k تعداد مدهای تابع متعامدتجربی استفاده شده برای بازسازی است.

مرحله سوم: اولین تجزیه SVD در داده‌های اصلی و جدید انجام می‌شود و یک ماتریس جدید توسط داده‌های موجود با مقادیر اصلی دوباره بازسازی شده و داده‌های از دست رفته توسط معادله بالا با $k=1$ محاسبه می‌شود. سپس تجزیه بعدی در ماتریس جدید بازسازی شده، تشکیل می‌شود و داده‌های از دست رفته دوباره محاسبه می‌شود. این فرآیند تا رسیدن به معیار همگرایی از پیش تعریف شده و موقعی که خطای میانگین مربع ریشه (RMSE) در نقاط اعتبارمتقابل تثبیت شود (تفاوت نسبی بین تکرارهای قبلی و فعلی کوچکتر از مقدار آستانه 1.0×10^5 است) تکرار می‌شود.

مرحله چهارم: تعداد تابع متعامدتجربی (k) های بازسازی شده با $k=2,3,\dots,k_{max}$ تکرار و روش ۳ نیز تکرار می‌شود. تعداد مطلوب توابع متعامدتجربی (k_{max}) حفظ می‌شود تا زمانی که حداقل RMSE بدست آید.

مرحله پنجم: پس از تعیین تعداد بهینه تابع متعامدتجربی، کل فرآیند از محصولات اعتبارسنجی که قبلاً کنار گذاشته شده بود، دوباره شروع می‌شود. سرانجام داده‌های کامل (فاقد شکاف) بازسازی شده با مقادیری که توسط داده‌های از دست رفته محاسبه شده با توابع متعامدتجربی کوتاه شده و مقادیر داده‌های موجود که برای مقادیر اصلی نگهداری می‌شوند، محاسبه می‌گردد. مرحله ششم: در نهایت محصولات تکمیل و نقاط فاقد داده پر می‌شوند.

ارزیابی صحت بازسازی به روش DINEOF

داده‌های میدانی در خلیج فارس و یا بخش‌هایی از آن و همچنین دریای عمان وجود دارد ولی این داده‌ها به صورت منظم و یکپارچه برای تمام سال‌های مورد مطالعه وجود نداشته است و حتی ممکن است دسترس به این داده‌ها به سهولت صورت نگیرد، اما دسترسی به داده‌های ماهواره‌ای راحت‌تر و همچنین بازسازی محصولات مفقوده در محیط R سریع‌تر صورت می‌گیرد. به منظور تأیید صحت هر بازسازی، چندین شاخص از جمله ضریب رگرسیون (R^2)، نسبت سیگنال به نویز (SNR)، RMSE، خطای مطلق میانگین (MAD) و میانگین مربع خطاها (MSE) استفاده شد. SNR به عنوان نسبت انحراف استاندارد مقادیر بازسازی شده به خطای استاندارد (تفاوت بین مقادیر اصلی و بازسازی شده) تعریف و شاخص‌های دیگر به شرح زیر محاسبه می‌شوند:

$$RMSE = \sqrt{\frac{\sum |S - I|^2}{N}} \quad MAD = \frac{\sum |S - I|}{N} \quad MSE = \frac{\sum |S - I|^2}{N}$$

در اینجا S شاخص مقادیر اصلی کلروفیل a (قبل از بازسازی)، I شاخص مقادیر بازسازی شده کلروفیل a و N تعداد کل داده‌ها می‌باشد [۱۸]. در مرحله پایانی نقشه‌های اصلی و بازسازی شده در نرم‌افزار ArcGIS 10.5 ایجاد و با هم مقایسه شدند.

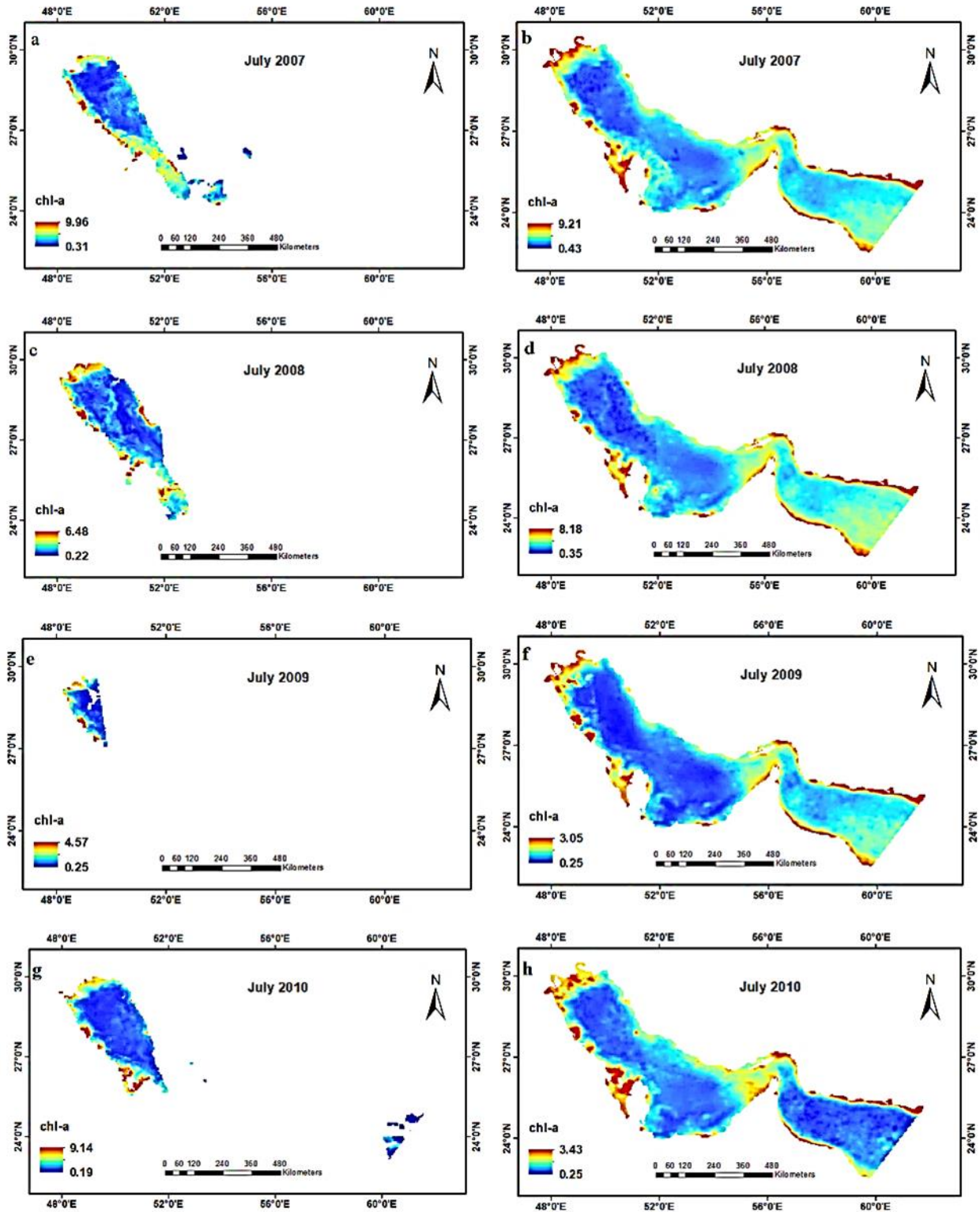
نتایج

در بررسی داده‌های استخراج شده مشخص شد که بیشترین تصاویر دارای محصولات مفقوده مربوط به ماه جولای و سال‌های ۲۰۰۷ تا ۲۰۱۸ بوده است. لذا در بخش نتایج، بر روی بازسازی این تصاویر تمرکز بیشتر شده است که در ادامه به گزارش خروجی‌های حاصل از بازسازی برای

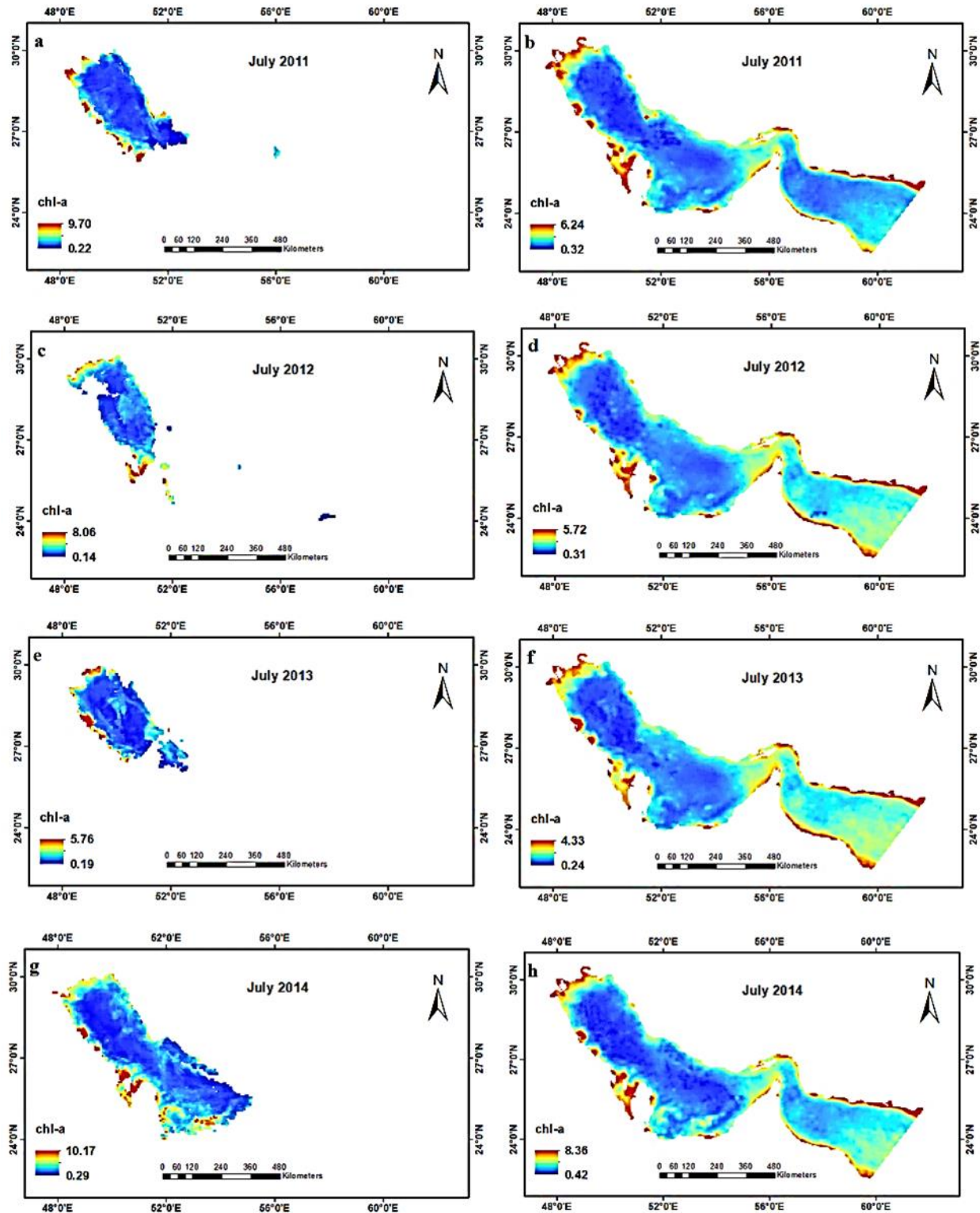
ماه جولای سال‌های ۲۰۰۷ تا ۲۰۱۸ اشاره می‌شود. البته برای بازسازی از کل محصولات غلظت کلروفیل a با وضوح مکانی ۴ کیلومتر حاصل از سنجنده مودیس آکوا (ژانویه ۲۰۰۳ تا دسامبر ۲۰۲۰) استفاده شده است. در این مطالعه مجموعه‌ای جدیدی (به شکل ماتریس) از داده‌های غلظت کلروفیل a با ۱۱۸۱۴۴ پیکسل مکانی و ۲۱۶ پیکسل زمانی با استفاده از روش DINEOF در نرم‌افزار R ایجاد شد. شکل‌های ۲ تا ۴ تصاویر دارای نقاط فاقد محصول و بازسازی شده‌ای غلظت کلروفیل a را در ماه جولای سال‌های ۲۰۰۷ تا ۲۰۱۸ در منطقه خلیج فارس و دریای عمان به نمایش گذاشته است. پس از بازسازی، ارزیابی صحت آن به وسیله محاسبه شاخص‌های $RMSE$ ، MAD ، MSE و SNR برای بازه زمانی جولای ۲۰۰۷ تا ۲۰۱۸ صورت گرفته است (جدول ۲). در ماه جولای سال‌های ۲۰۰۹ ($RMSE=0.056$ و $RMSE=0.053$) و $SNR=0.70$) ۲۰۱۵ و $SNR=0.49$ و بیشترین میزان نقاط فاقد داده در سال‌های مورد مطالعه دیده می‌شود و همچنین تصاویر غلظت کلروفیل a دریای عمان نسبت به خلیج فارس دارای نقاط فاقد محصول بیشتری بوده است. نتایج بررسی مقدار شاخص‌های ارزیابی صحت (SNR ، MAD ، MSE و $RMSE$) بازسازی محصولات مفقوده به وسیله الگوریتم DINEOF که برای بازه زمانی جولای ۲۰۰۳ تا ۲۰۱۸ در منطقه خلیج فارس و دریای عمان در این مطالعه صورت گرفته است، نشان می‌دهند که کمترین و بیشترین $RMSE$ مربوط به جولای ۲۰۰۷ (0.32) و 0.15 (0.70)، کمترین و بیشترین مقدار MAD مربوط به جولای ۲۰۰۷ (0.05) و 0.15 (0.32)، کمترین و بیشترین مقدار MSE مربوط به جولای ۲۰۱۳ (0.08) و 0.15 (0.49)، کمترین و بیشترین مقدار R^2 مربوط به جولای ۲۰۱۵ (0.49) و 0.14 (0.83) و در نهایت کمترین و بیشترین مقدار SNR مربوط به جولای ۲۰۰۹ (1.14) و 2.08 (0.08) بوده است. شکل‌های ۲ تا ۴ نشان می‌دهد که در نواحی ساحلی میزان غلظت کلروفیل a نسبت به نواحی مرکزی و شمال غربی منطقه مورد مطالعه بیشتر و در بیشتر سال‌های مورد مطالعه مقدار غلظت کلروفیل a در دریای عمان بیشتر از خلیج فارس بوده است. بیشترین مقدار غلظت کلروفیل a در بازه زمانی جولای ۲۰۱۸-۲۰۰۷ مربوط به سال ۲۰۰۷ ($9.21-0.43$ میلیگرم بر متر مکعب) و کمترین مقدار مربوط به سال ۲۰۰۹ ($3.04-0.2$ میلیگرم بر متر مکعب) بوده است.

جدول ۱. مقدار پارامترهای $RMSE$ ، MAD ، MSE ، R^2 و SNR در بازه زمانی جولای ۲۰۰۷ تا ۲۰۱۸ در خلیج فارس و دریای عمان

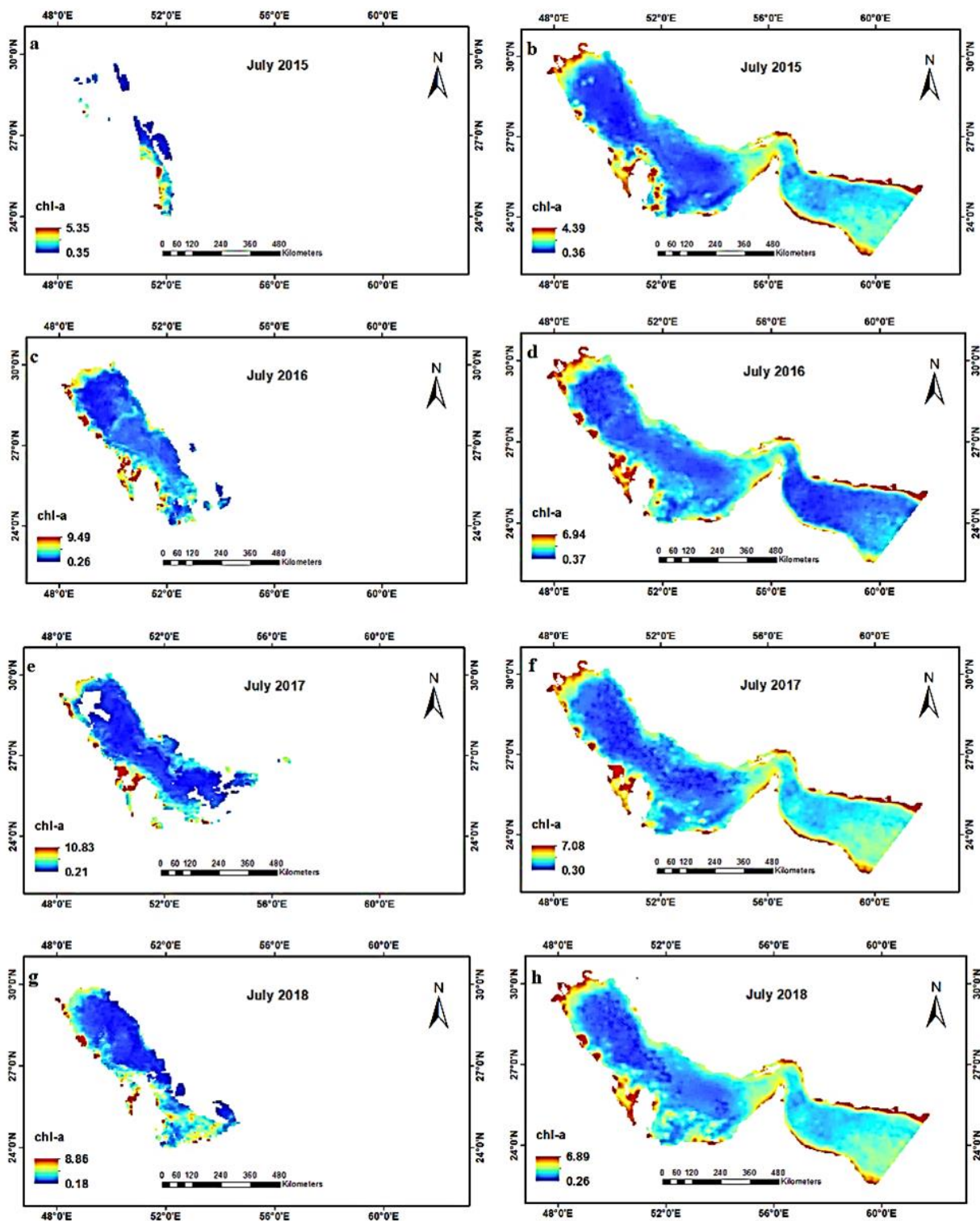
Year	RMSE	MAD	MSE	R ²	SNR
۲۰۰۷	۰/۴۵	۰/۰۵	۰/۲۰	۰/۷۴	۱/۵۹
۲۰۰۸	۰/۴۲	۰/۱۹	۰/۱۷	۰/۸۰	۱/۸۱
۲۰۰۹	۰/۵۶	۰/۲۷	۰/۳۲	۰/۵۳	۱/۱۴
۲۰۱۰	۰/۶۲	۰/۲۱	۰/۳۸	۰/۶۲	۱/۲۶
۲۰۱۱	۰/۳۲	۰/۱۵	۰/۱۰	۰/۷۶	۱/۷۰
۲۰۱۲	۰/۳۵	۰/۱۴	۰/۱۲	۰/۷۹	۱/۷۳
۲۰۱۳	۰/۲۹	۰/۱۰	۰/۰۸	۰/۷۵	۱/۴۸
۲۰۱۴	۰/۳۲	۰/۱۵	۰/۱۰	۰/۸۳	۲/۰۸
۲۰۱۵	۰/۷۰	۰/۳۲	۰/۴۹	۰/۴۹	۱/۴۷
۲۰۱۶	۰/۳۶	۰/۱۶	۰/۱۳	۰/۷۷	۱/۸۲
۲۰۱۷	۰/۴۵	۰/۲۱	۰/۲۰	۰/۷۴	۱/۶۰
۲۰۱۸	۰/۴۶	۰/۲۲	۰/۲۱	۰/۷۰	۱/۴۸



شکل ۲. شکل‌های a-g نشان دهنده نقشه‌های اصلی و دارای گسست داده‌های غلظت کلروفیل *a* (میلیگرم بر متر مکعب) و شکل‌های b-h نشان دهنده نقشه‌های بازسازی شده آن در سال‌های ۲۰۱۰-۲۰۰۷ در خلیج فارس و دریای عمان



شکل ۳. شکل‌های a-g نشان دهنده نقشه‌های اصلی و دارای گسست داده‌ای غلظت کلروفیل *a* (میلیگرم بر متر مکعب) و شکل‌های b-h نشان دهنده نقشه‌های بازسازی شده آن در سال‌های ۲۰۱۱-۲۰۱۴ در خلیج فارس و دریای عمان



شکل ۴. شکل‌های a-g نشان دهنده نقشه‌های اصلی و دارای گسست داده‌ای غلظت کلروفیل *a* (میلیگرم بر متر مکعب) و شکل‌های h نشان دهنده نقشه‌های بازسازی شده آن در سال‌های ۲۰۱۵-۲۰۱۸ در خلیج فارس و دریای عمان

بحث

در مطالعه‌ای از ۱۰ الگوریتم که شامل DINEOF، درون‌یابی‌های زمین‌آمار و روش‌های یادگیری تحت‌نظارت می‌باشند، در دریای بوفورت جنوبی، دریایی چوکچی جنوبی، بخشی از مناطق گرمسیری و خلیج مکزیک برای بازسازی محصول کلروفیل *a* استفاده شد. نتایج این تحقیق نشان می‌دهد که در بین الگوریتم‌های استفاده شده، سه الگوریتم DINEOF، کریجینگ معمولی و کریجینگ مکانی و زمانی دارای دقت بالاتر از متوسط در همه زمینه‌ها و وظایف مورد مطالعه مثل بررسی صحت بازسازی بودند. این الگوریتم‌ها در بسته‌های نرم‌افزاری مختلف در دسترس هستند. کریجینگ معمولی نیاز به تلاش دستی و زمان محاسباتی متوسط داشت. DINEOF خطاهای بزرگتری داشت و همچنین زمان محاسباتی متوسط برای آن لازم بود، اما هنوز هم به طور مداوم خوب کار می‌کند و نیازی به محاسباتی مثل محاسبه واریوگرام ندارد. علاوه بر این DINEOF آزمایش شده است و در بسیاری از تنظیمات دیگر عملکرد خوبی دارد [۱۷، ۱۹، ۳۷-۳۴]. برعکس، کریجینگ مکانی-زمانی یکی از وقت‌گیرترین الگوریتم‌ها چه از نظر تلاش دستی و چه از نظر زمان محاسباتی بود که آزمایش شد [۱۷]. در تحقیقی از الگوریتم DINEOF برای بازسازی محصولات مفقوده کلروفیل *a* خلیج اسپنگار استفاده شده است. در این مطالعه نیز مدل DINEOF برای اهداف اعتبارسنجی با مدل کریجینگ مکانی زمانی مقایسه شدند. نتایج بدست آمده نشان می‌دهد که مدل DINEOF دارای بیشترین ضریب همبستگی پیرسون (۰/۹۹۴۰) و کوچکترین مقادیر RMSE (۰/۲۷۷)، MAD (۰/۱۵۵) در منطقه مورد مطالعه است. از این رو این مطالعه نشان داد که مدل DINEOF برای پر کردن موقعیت‌های مکانی طولانی فاقد داده مؤثرتر است [۳۸]. بدلیل مزیت‌های مطرح شده که در مطالعات اشاره شده و همچنین مزیت‌های از قبیل استفاده در ماتریس‌های بزرگ، قوی و ساده بودن روش، عدم نیاز به اطلاعات قبلی در مورد آمار خطای داده‌ها در مطالعه حاضر برای بازسازی داده‌های سطح ۳ (۴ کیلومتر) مودیس ۲۰۰۳ تا ۲۰۲۰ کلروفیل *a* در منطقه خلیج فارس و دریای عمان از الگوریتم DINEOF در نرم‌افزار R استفاده شد. بدلیل وجود محصولات مفقوده بیشتر در ماه جولای و سال‌های ۲۰۰۷ تا ۲۰۱۸، از این بازه زمانی در این مطالعه استفاده شده است. نرم‌افزار R یک نرم‌افزار با دسترسی آزاد و رایگان بوده و دارای سرعت پردازش بالا است و در این روش نیازی به استفاده از پردازشگرهای قوی و پیچیده نیست و همچنین مستقیماً روی داده‌های دانلودی کار می‌کند. در مطالعه‌ای محققان برای اولین بار الگوریتم DINEOF را برای پر کردن داده‌های از دست رفته توصیف کردند و این روش را برای داده‌های دمایی آب سطح دریا آدریاتیک به کار بردند. در این مطالعه به معرفی الگوریتم DINEOF برای استنباط داده‌های مفقوده از سری داده‌های اقیانوس‌شناسی، تعیین EOF بهینه با روش اعتبارسنجی متقابل برای مجموعه داده‌های ناقص و کامل و همچنین سطح نویز و ارزیابی صحت‌بازسازی به وسیله تعیین خطاهای درون‌یابی پرداخته شده است. در این مطالعه تعداد EOF بهینه ۴ بدست آمده است. در بخشی از این مطالعه با به کارگیری روش درون‌یابی خطی مکانی و بدون در نظر گرفتن بُعد زمانی به مقایسه آن با الگوریتم DINEOF پرداخته شده است. نتایج این مقایسه نشان داده است که در روش درون‌یابی خطی مکانی خطا از ۰/۰۶۷ به ۰/۰۲۴ و در الگوریتم DINEOF از ۰/۰۶۴ به ۰/۰۰۲ کاهش یافته است. نتایج این مطالعه همچنین نشان داد که در روش جدید امکان تعیین تعداد EOF بهینه و آنالیز کمی خطا وجود دارد [۱۷]. در مطالعه‌ای دیگر از الگوریتم DINEOF برای بازسازی داده‌های مفقوده برای مجموعه داده‌های بزرگ برای داده‌های SST (Sea Surface Temperature) آدریاتیک استفاده کردند. این روش با مقادیر مختلف داده‌های مفقود آزمایش شده است، که به طور مصنوعی ابرهایی از رنج ۴۰ تا ۸۰ درصد از دست دادگی داده را اضافه کرده و سپس با مجموعه داده‌های مشابه (با داده‌هایی که از دست نرفته‌اند) و همچنین با داده‌های درمکان مقایسه شده‌اند. مطالعات اعتبارسنجی نشان می‌دهد که حتی زمانی که مقدار داده‌های مفقود شده بسیار زیاد باشد، نتایج قوی هستند. بازسازی داده‌های دریای آدریاتیک ویژگی‌های واقع‌بینانه و توزیع قابل اعتماد دما را نشان می‌دهد. اعتبارسنجی در یک مجموعه داده با افزایش مقدار داده‌های مفقوده انجام شد. مقایسه داده‌های بازسازی شده و اصلی نشان می‌دهد که خطا کوچک است و به ترتیب حدود ۰/۸۹، ۰/۷۸ و ۱/۲۵ درجه سانتیگراد برای ۴۰، ۶۰ و ۸۰ درصد از داده‌های مفقوده است. مقایسه با داده‌های درمکان از پایگاه داده MEDAR/Medatlas همچنین یک خطای RMS با همان اندازه را که در حدود ۰/۹۵ درجه سانتیگراد برای نقاطی که در مجموعه کامل توسط ابرها پوشانده شده‌اند را نشان می‌دهد [۱۷]. میزان خطا در مطالعه حاضر در ماه‌های ۲۰۰۹ و ۲۰۱۵ که دارای

داده‌های مفقوده زیادی هستند به ترتیب ۰/۵۶ و ۰/۷۰ بوده است که نسبت به این مطالعه دارای میزان خطای کمتری می‌باشد. در این مطالعه برای ارزیابی صحت‌بازسازی محصولات کلروفیل *a* در خلیج فارس و دریای عمان پارامترهای *RMSE*، *MAD*، *MSE*، *R*² و *SNR* محاسبه شده است. نتایج نشان می‌دهد که در ماه جولای ۲۰۰۹ (*SNR* = ۱/۱۴، *R*² = ۰/۵۳، *MSE* = ۰/۳۲، *MAD* = ۰/۲۷، *RMSE* = ۰/۵۶) و ۲۰۱۵ (*SNR* = ۱/۴۷، *R*² = ۰/۴۹، *MSE* = ۰/۴۹، *MAD* = ۰/۳۲، *RMSE* = ۰/۷۰) علت میزان خطای بالا بین محصولات اصلی و بازسازی شده در این زمان‌ها باشد. در جولای ۲۰۱۴ (*MAD* = ۰/۱۵، *RMSE* = ۰/۳۲) تا ۲۰۱۴ از نظر صحت بازسازی با در نظر گرفتن پارامترهای اشاره شده بهترین عملکرد را از نظر بازسازی داشته است. به طور کلی در بازه زمانی ۲۰۱۱ تا ۲۰۱۴ از نظر صحت بررسی میزان پارامترهای ذکر شده در بالا می‌توان این چنین نتیجه گرفت که استفاده از الگوریتم DINEOF در نرم‌افزار R نتایج قابل قبولی را ارائه می‌دهد، اما در برخی از سال‌ها به طور مثال ۲۰۱۵ ممکن است نتایج ضعیفی ارائه دهد که می‌تواند علل مختلفی داشته باشد. هر روشی به نوبه خود می‌تواند دارای نقاط ضعف و قوتی باشند که برای برطرف کردن نقاط ضعف روشی مثل DINEOF می‌توان از الگوریتم‌های اصلاح شده آن و یا از زبان‌های برنامه‌نویسی مثل پایتون و نرم‌افزارهای مانند QGIS استفاده کرد.

نتیجه‌گیری نهایی

به طور کلی می‌توان این گونه نتیجه گرفت که بهترین الگوریتم بازسازی محصولات مفقوده بستگی به منطقه مورد مطالعه، داده‌ها و هدف از پر کردن گسست داده‌ای و دوره پوشش به وسیله تصاویر دارد. کریجینگ معمولی، کریجینگ مکانی-زمانی و DINEOF در مناطق مورد مطالعه و وظایف مختلف به خوبی کار کرده‌اند. بسته‌های نرم‌افزاری مختلف در دسترس هستند و بنابراین به عنوان روش‌های عمومی بازسازی داده‌های ماهواره‌ای دریایی مفقوده برای آنها توصیه می‌شود. برای انتخاب بین این الگوریتم‌ها باید به تفکیک زمانی داده‌ها توجه کرد. در مطالعات آینده می‌توان از روش‌های دیگر بازسازی داده‌های ماهواره‌ای مثل ماشین‌یادگیری، شبکه‌عصبی و همچنین نرم‌افزار DINEOF و (variable VE- DINEOF (EOFs- DINEOF که حالت بهینه تابع متعامدتجربی را به صورت محلی پیدا می‌کند و دقت بازسازی را بالا می‌برد، در این منطقه استفاده کرد و با نتایج این مطالعه مورد مقایسه قرار داد. با مقایسه الگوریتم‌های مختلف بازسازی محصولات مفقوده می‌توان بهترین الگوریتم را برای این منطقه پیدا کرد. همچنین می‌توان از محصولات در بازه زمانی طولانی‌تر و مکانی وسیع‌تر از این روش استفاده کرد و به ارزیابی صحت‌بازسازی آن پرداخت. محصولات کلروفیل *a* حاصل از بازسازی برای کاربردهایی مانند تشخیص شکوفایی جلبکی و همچنین مدیریت مناسب شیلاتی قابل استفاده است. در مقیاس زمانی بلندمدت، مقایسه محصولات مفقوده با بازسازی شده می‌تواند به کنترل کیفیت داده‌های ماهواره‌ای کمک می‌کند.

تأییدیه اخلاقی: مورد خاصی توسط نویسندگان گزارش نشده است.

تعارض منافع: مورد خاصی توسط نویسندگان گزارش نشده است.

منابع

- 1- Khan IA, Ghazal L, Arsalan MH, Siddiqui MF, Kazmi JH. Assessing spatial and temporal variability in phytoplankton concentration through chlorophyll-a satellite data: A case study of northern Arabian Sea. *Pak J Bot.* 2015;47(2):797-805.
- 2- Hattab T, Jamet C, Sammari C, Lahbib S. Validation of chlorophyll- α concentration maps from Aqua MODIS over the Gulf of Gabes (Tunisia): Comparison between MedOC3 and OC3M bio-optical algorithms. *International Journal of Remote Sensing.* 2013;34(20):7163-77.

- 3- McClain CR. A decade of satellite ocean color observations. *Annual Review of Marine Science*. 2009; 1:19-42.
- 4- Chauhan P, Nagur C, Mohan M, Nayak S, Navalgund R. Surface chlorophyll-a distribution in Arabian Sea and Bay of Bengal using IRS-P4 Ocean Colour Monitor satellite data. *Current Science*. 2001;80(2):127-129.
- 5- Radiarta IN, Saitoh S-I. Satellite-derived measurements of spatial and temporal chlorophyll-a variability in Funka Bay, southwestern Hokkaido, Japan. *Estuarine, Coastal and Shelf Science*. 2008;79(3):400-408.
- 6- Behrenfeld MJ, Falkowski PG. Photosynthetic rates derived from satellite-based chlorophyll concentration. *Journal of Limnology and oceanography*. 1997;42(1):1-20.
- 7- Alvera-Azcárate A, Barth A, Beckers JM, Weisberg RH. Multivariate reconstruction of missing data in sea surface temperature, chlorophyll, and wind satellite fields. *Journal of Geophysical Research: Oceans*. 2007;112(C3).
- 8- Saulquin B, Gohin F, Garrello R. Regional objective analysis for merging high-resolution MERIS, MODIS/Aqua, and SeaWiFS chlorophyll-a data from 1998 to 2008 on the European Atlantic shelf. *IEEE transactions on geoscience and remote sensing*. 2010;49(1):143-154.
- 9- Pukhtyar L, Stanichny S, Timchenko I. Optimal interpolation of the data of remote sensing of the sea surface. *Physical Oceanography*. 2009;19(4):225-239.
- 10- Bennett AF. *Inverse modeling of the ocean and atmosphere*: Cambridge University Press; 2005.
- 11- Kaplan A, Kushnir Y, Cane MA, Blumenthal MB. Reduced space optimal analysis for historical data sets: 136 years of Atlantic sea surface temperatures. *Journal of Geophysical Research: Oceans*. 1997;102(C13):27835-27860.
- 12- Beckers J-M, Rixen M. EOF calculations and data filling from incomplete oceanographic datasets. *Journal of Atmospheric and Oceanic Technology*. 2003;20(12):1839-1856.
- 13- Alvera-Azcárate A, Barth A, Rixen M, Beckers J-M. Reconstruction of incomplete oceanographic data sets using empirical orthogonal functions: application to the Adriatic Sea surface temperature. *Ocean Modelling*. 2005;9(4):325-346.
- 14- Alvera-Azcárate A, Barth A, Sirjacobs D, Lenartz F, Beckers J-M. Data Interpolating Empirical Orthogonal Functions (DINEOF): a tool for geophysical data analyses. *Mediterranean Marine Science*. 2011;12(3):5-11.
- 15- Beckers J-M, Barth A, Alvera-Azcárate A. DINEOF reconstruction of clouded images including error maps—application to the Sea-Surface Temperature around Corsican Island. *Ocean Science*. 2006;2(2):183-199.
- 16- Lehoucq RB, Sorensen DC, Yang C. *ARPACK users' guide: solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods*: SIAM; 1998.
- 17- Stock A, Subramaniam A, Van Dijken GL, Wedding LM, Arrigo KR, Mills MM, et al. Comparison of cloud-filling algorithms for marine satellite data. *Remote Sensing*. 2020;12(20):3313.
- 18- Wang Y, Liu D. Reconstruction of satellite chlorophyll-a data using a modified DINEOF method: a case study in the Bohai and Yellow seas, China. *International Journal of Remote Sensing*. 2014;35(1):204-217.
- 19- Hilborn A, Costa M. Applications of DINEOF to satellite-derived chlorophyll-a from a productive coastal region. *Remote Sensing*. 2018;10(9):1449.
- 20- Ji C, Zhang Y, Cheng Q, Tsou J, Jiang T, San Liang X. Evaluating the impact of sea surface temperature (SST) on spatial distribution of chlorophyll-a concentration in the East China Sea. *International Journal of Applied Earth Observation and Geoinformation*. 2018; 68:252-261.

- 21- Moradi M, Moradi N. Correlation between concentrations of chlorophyll-a and satellite derived climatic factors in the Persian Gulf. *Marine Pollution Bulletin*. 2020;1 61:111728.
- 22- Zhou Y, Gai Y, Li J, editors. *Research on Sea Surface temperature Reconstruction from long-term MODIS data*. IOP Conference Series: Earth and Environmental Science; 2021: IOP Publishing.
- 23- Valentini E, Filippini F, Nguyen Xuan A (ISPRA); Dontas I, Bogonikolos N (ARATOS); Manakos I, Kordelas G, Marini K, Kotsias V, Papanastasiou M (CERTH); Bustamante J, Díaz Delgado R (CSIC); Karnieli A, Ohana-Levi N, Stein A (BGU); Tarantino C, Blonda P (CNR-IIA). *Ecopotential: Improving future ecosystem benefits through earth observations*; 2017. 13.
- 24- Rossi RE, Dungan JL, Beck LR. Kriging in the shadows: geostatistical interpolation for remote sensing. *Remote Sensing of Environment*. 1994;49(1):32-40.
- 25- Vaughan GO, Al-Mansoori N, Burt JA. *The arabian gulf. World seas: An environmental evaluation*: Elsevier; 2019. p. 1-23.
- 26- Grafton RQ, Hilborn R, Squires D, Williams M, Tait M. *Handbook of marine fisheries conservation and management*: OUP USA; 2010.
- 27- Jawad LA. *Dangerous Fishes of the Eastern and Southern Arabian Peninsula*. 2019.
- 28- Chaichitehrani N, Allahdadi MN. Overview of wind climatology for the Gulf of Oman and the northern Arabian Sea. *American Journal of Fluid Dynamics*. 2018;8(1):1-9.
- 29- Nikoian Aaa. *Study of hydrology and hydrobiology of the Persian Gulf in the waters of Khuzestan, Bushehr and Hormozgan*. Iranian Fisheries Research Institute. Tehran.1991; 116 p. (in persian).
- 30- ROPME. *State of the Marine Environment Report 2003*. Regional Organization for the Protection of the Marine Environment. Kuwait. 2004: 217.
- 31- Sheppard C, Price A, Roberts C. *Marine ecology of the Arabian region: patterns and processes in extreme tropical environments*.1992.
- 32- Aicken W, Shimwell S, Stapleton N. Improved monitoring of oceanographic features in the Gulf of Oman through combined use of satellite thermal infra-red, ocean colour & radar altimeter observations. European Space Agency, (Special Publication) ESA SP. 2000; 218-224.
- 33- Wiggert J, Murtugudde R, McClain C. Processes controlling interannual variations in wintertime (Northeast Monsoon) primary productivity in the central Arabian Sea. *Deep Sea Research Part II: Topical Studies in Oceanography*. 2002;49(12):2319-2343.
- 34- Alvera-Azcárate A, Vanhellemont Q, Ruddick K, Barth A, Beckers J-M. Analysis of high frequency geostationary ocean colour data using DINEOF. *Estuarine, Coastal and Shelf Science*. 2015; 159:28-36.
- 35- Sirjacobs D, Alvera-Azcárate A, Barth A, Lacroix G, Park Y, Nechad B, et al. Cloud filling of ocean colour and sea surface temperature remote sensing products over the Southern North Sea by the Data Interpolating Empirical Orthogonal Functions methodology. *Journal of Sea Research*. 2011;65(1):114-30.
- 36- Li Y, He R. Spatial and temporal variability of SST and ocean color in the Gulf of Maine based on cloud-free SST and chlorophyll reconstructions in 2003–2012. *Remote Sensing of Environment*. 2014; 144:98-108.
- 37- Liu X, Wang M. Gap filling of missing data for VIIRS global ocean color products using the DINEOF method. *IEEE Transactions on Geoscience and Remote Sensing*. 2018;56(8):4464-76.
- 38- Yussof F, Maan NB, Reba M. Reconstruction of Chlorophyll-a Data by Using DINEOF Approach in Sepanggar Bay, Malaysia. *Computer Science*. 2021; 16:345-56.

Reconstruction of MODIS chlorophyll *a* products using DINEOF algorithm in R software: A case study of the Persian Gulf and Oman Sea

Maryam Karimian¹, Omid Beyraghdar Kashkooli^{1*}, Reza Modarres¹, Saeid Pourmanafi¹

1. Department of Natural Resources, Isfahan University of Technology, Isfahan, Iran

ABSTRACT

The purpose of this study was to use the DINEOF algorithm to fill in the missing data points in the satellite imagery of chlorophyll *a* in the surface waters of the Persian Gulf and Oman Sea. In this study, level 3 data (spatial resolution of 4 km) of chlorophyll *a* concentration obtained from MODIS sensor in the period 2003 to 2020 in the Persian Gulf and Oman Sea (extracted from NASA site) were used. Some images had no product points in some months of the year. Images with missing products (gaps) of chlorophyll *a* concentration in the Persian Gulf and Oman Sea were reconstructed using `rtsa.gapfill` software package and DINEOF algorithm in R software. Finally, to evaluate the accuracy of the DINEOF algorithm, parameters such as RMSE, MSE, MAD and SNR have been calculated. The original maps (with gaps) of chlorophyll *a* concentration showed that highest amount of lost product was in July 2009 and 2015. Assessing the accuracy of the reconstruction, it was found that the reconstruction of the missing data was done well by the algorithm (for example, in July 2014, the calculated parameters are as follows: $R^2 = 0.83$, $RSME = 0.34$, $MAD = 0.14$ and $OMSE = 0.10$). The results showed that the implementation of DINEOF algorithm in R software to reconstruct the missing products in chlorophyll *a* images is an efficient, rapid and successful method. Reconstructed chlorophyll *a* products can be used for applications such as detecting algal blooms as well as proper fisheries management. On a long-term scale, comparing lost products with reconstructed ones helps control the quality of satellite data.

KEYWORDS: Chlorophyll-*a* concentration, Products reconstruction, DINEOF, The Persian Gulf, Oman Sea

ARTICLE TYPE

Original Research

ARTICLE HISTORY

Received: 23 Sept 2021

Accepted: 21 April 2022

ePublished: 22 May 2022

* Corresponding Author:

Email address: omid.beyraghdar@iut.ac.ir

Tel: +(98)

© Published by Tarbiat Modares University

eISSN:2476-6887 pISSN:2322-5513