



Survey of Changes in Colored Dissolved Organic Matter (CDOM) in Surface Water Resources Using SVM Algorithm and Landsat 8 Satellite Images

M. Momeni Esfahani^{1*} and A.S.H. Amini²

Abstract

Colored dissolved organic matter (CDOM), one of the main constituents of DOM in surface water, is an important indicator of water quality, the biochemical status, and the nutritious material content and plays an important role in the carbon cycle in surface water resources. In this study, we have analyzed colored dissolved organic matter (CDOM) in lakes at the watershed area of North Siberia for the period of 2013 to 2016 using Landsat 8 images. The support vector regression model was used for selecting the most desirable band in determining the CDOM absorption coefficient. Using the support vector machine model to classify and compare the changes in the amount of $\alpha_{\text{CDOM}}(440)$, the CDOM scatter plots for the years of 2014 and 2015 were obtained. Based on the results, regarding the correlation coefficient ($R^2=0/71$) and the amount of errors ($\text{MSE}=1/60 \text{ m}^{-1}$, $\text{RMSE}=1/0775 \text{ m}^{-1}$, and $\text{MAE}=0/9464 \text{ m}^{-1}$), it was concluded that the application of green/red band ratio in Landsat 8 satellite was the most desirable choice for measuring CDOM at a wavelength of 440 nm in watershed resources of North Russia. The scatter plot indicated an increase in CDOM in the lakes in the northeast of the area in 2015 compared to 2014.

Keywords: Colored Dissolved Organic Matter, SVR, SVM, Landsat 8.

Received: September 5, 2020

Accepted: December 13, 2020

بررسی تغییرات ماده آلی محلول رنگی (CDOM) با استفاده از الگوریتم SVM و تصاویر ماهواره لندست ۸ در منابع آبهای سطحی

محمد مومنی اصفهانی^{۱*} و امیر شاهرخ امینی^۲

چکیده

ماده آلی محلول رنگی (CDOM) یکی از اجزای اصلی α_{CDOM} در آبهای سطحی است و شاخصی مهم در کیفیت آب، وضعیت بیوشیمی و محتوای مواد مغذی است و نقش مهمی در چرخه کربن در آبهای سطحی دارد. در این پژوهش به تحلیل ماده آلی محلول رنگی (CDOM) در دریاچه‌ها با استفاده از تصاویر لندست ۸ در سال‌های ۲۰۱۳ تا ۲۰۱۶ در دریاچه‌های شمال سیبری پرداخته شده است. از مدل بهینه رگرسیون بردار پشتیبان (GA-SVR) برای انتخاب مناسب‌ترین باند در تعیین ضریب جذب CDOM استفاده شده است و با استفاده از مدل ماشین بردار پشتیبان (SVM) به منظور طبقه‌بندی و مقایسه تغییرات میزان $\alpha_{\text{CDOM}}(440)$ ضریب جذب CDOM در طیف ۴۴۰ نانومتر نقشه مکانی پراکندگی CDOM در سال‌های ۲۰۱۴ و ۲۰۱۵ بدست آمده است. بر اساس نتایج حاصله، با توجه به ضریب همبستگی ($R^2=0/71$) و میزان خطاها ($\text{MSE}=1/60 \text{ m}^{-1}$ ، $\text{RMSE}=1/0775 \text{ m}^{-1}$ ، $\text{MAE}=0/9464 \text{ m}^{-1}$) نتیجه گرفته شد که استفاده از نسبت باندهای سبز به قرمز در ماهواره لندست ۸ مناسب‌ترین انتخاب برای تعیین ماده آلی محلول رنگی در طول موج ۴۴۰ نانومتر در منطقه مطالعاتی است. نقشه پراکندگی نمایانگر افزایش مقدار ماده آلی محلول رنگی در دریاچه‌های شمال-شرقی منطقه در سال ۲۰۱۵ نسبت به سال ۲۰۱۴ است.

کلمات کلیدی: ماده آلی محلول رنگی، رگرسیون بردار پشتیبان، ماشین بردار پشتیبان، لندست ۸.

تاریخ دریافت مقاله: ۹۹/۶/۱۵

تاریخ پذیرش مقاله: ۹۹/۹/۲۳

1- M.Sc. Graduate of Remote Sensing, Department of Surveying Engineering, Islamic Azad University, South Tehran Branch, Tehran, Iran Email: salar.me3@gmail.com

2- Associate Professor, Department of Surveying Engineering, Islamic Azad University South Tehran Branch, Tehran, Iran. Email: Sh_amini@azad.ac.ir

*- Corresponding Author

۱- دانش‌آموخته کارشناسی ارشد سنجش از دور، گروه مهندسی نقشه‌برداری، دانشگاه آزاد اسلامی، واحد تهران جنوب، تهران، ایران.

۲- استادیار گروه مهندسی نقشه‌برداری، دانشگاه آزاد، واحد تهران جنوب، تهران، ایران.

*- نویسنده مسئول

بحث و مناظره (Discussion) در مورد این مقاله تا پایان زمستان ۱۳۹۹ امکانپذیر است.

۱- مقدمه

دریاچه‌ها بهترین منبع آب شیرین موجود در سطح زمین هستند و محیط مناسبی را برای انواع مختلف گونه‌های زیستی فراهم می‌کنند. در واقع دریاچه‌ها اجزای اصلی و ویژه‌ای از چرخه هیدرولوژی، مواد مغذی و کربن را تشکیل می‌دهند (Moss, 2012). عوامل مختلفی بر روی اکوسیستم آب‌های سطحی تأثیرگذار است که از DOM (ماده آلی محلول) می‌توان به عنوان مهمترین عامل نام برد. ثابت شده است که ماده آلی محلول (DOM) نقش مهمی در فرآیندهای بیوشیمیایی دارد: در واقع انرژی را به موجودات هتروتروفی منتقل و نفوذ نور را تحت تأثیر قرار می‌دهد؛ ضمن آنکه بر تشکیل تجمع ذرات و تعامل آن با ترکیبات آلی مصنوعی و فلزات تأثیر می‌گذارد (Jansson, 1992; Tranvik, 1998). منابع اصلی DOM که به آب‌های سطحی وارد می‌شود شامل: تخلیه رودخانه از مواد آلی حاصل از زمین و مواد آلی محلول ناشی از تجزیه شکوفه‌های فیتوپلانکتون است. در داخل منابع DOM، زیرمجموعه‌ای از مواد آلی که قادر به جذب بسیار زیاد فوتون‌های انرژی نور است (به ویژه اشعه ماوراء بنفش با انرژی بالا)، وجود دارد و تحت تأثیر انواع واکنش‌های شیمیایی ناشی از آزادسازی یون‌ها است که به این دسته از مواد جذب‌کننده نور، ماده آلی محلول رنگی (CDOM) گفته می‌شود. ساختار شیمیایی CDOM بسته به منبع اصلی آن و وضعیت موجود در تغییرات بیولوژیکی و فتوشیمیایی متفاوت است، اما ویژگی‌های مشترک آن شامل فنولیک^۳ جذب نور (Opsahl and Benner, 1998; Vähätalo et al., 1999; Opsahl and Zepp, 2001) و ویژگی‌های کربوکسیلیک است (Budac and Wan, 1992; Johannessen and Miller, 2001). اهمیت CDOM به دلیل جذب نور و کاهش نفوذ نور در طبقه‌های آب می‌باشد که باعث مختل شدن فعالیت‌های بیولوژیکی مانند فتوسنتز می‌شود، بنابراین از تولید فیتوپلانکتون‌ها که برای زنجیره غذایی ضروری می‌باشد، جلوگیری می‌کند (Häder et al., 2007).

در حال حاضر اکثر برنامه‌های نظارتی CDOM به صورت میدانی است که به علت نبود فرمول و یا شاخص کلی به منظور بازیابی این ماده به وسیله تکنیک‌های سنجش از دور می‌باشد. به طور معمول، برای اندازه‌گیری CDOM از روش‌های میدانی و آزمایشگاهی استفاده می‌شود که به صورت ضریب جذب (واحد آن m^{-1}) یا واحد RFU که با استفاده فلورسنس به دست آید، بیان می‌شود و هزینه و زمان زیادی را در بر می‌گیرد.

CDOM از لحاظ اپتیکی قابل اندازه‌گیری است و بنابراین کاندیدای عالی برای اندازه‌گیری با تکنیک‌های سنجش از دور است. در ابتدا

غالباً در داده‌های سنجش از دور به عنوان بخشی از مدل‌های بازتابی سطح، در هنگام تخمین کلروفیل a و غلظت رسوب معلق (SPM) در آب‌های سطحی برآورد می‌شود. محققان علاقه‌مند به اندازه‌گیری با استفاده از فناوری سنجش از دور CDOM در دهه ۱۹۸۰ به این نتیجه رسیده‌اند که قبل از اندازه‌گیری فیتوپلانکتون، کلروفیل a و رسوب معلق (SPM)، باید تأثیر CDOM از بازتاب داده‌های سنجش از دور به دست آید (Ferrari and Tassan, 1992; Karabashev et al., 1993; Tassan, 1988). با این حال، در برخی از مطالعات اخیر نیز تلاش شده است CDOM آب‌های سطحی از ماهواره‌ای چندطیفی سنجش از دور محاسبه شود. اما یکی از چالش‌های بزرگ موجود، فقدان روشی استاندارد برای تولید داده‌های اصلاح شده اتمسفری است (Kutser et al., 2009; Kutser et al., 2016).

در دهه‌های گذشته برای مطالعه توزیع و فراوانی CDOM اندازه‌گیری ضریب جذب CDOM از شیب طیفی این ماده از ۴۰۰ تا ۵۰۰ نانومتر در طیف الکترومغناطیسی استفاده شده است (Babin et al., 2003; Siegel et al., 2002; Nelson and Siegel, 2013; Organelli et al., 2014). از تلاش‌های انجام شده می‌توان به تلاش‌های (Kutser et al., 2015; 2016) نام برد که در سال ۲۰۰۵ فرمولی به منظور به دست آوردن ضریب جذب CDOM با استفاده از تصاویر ماهواره ALI به انجام رساندند که فرمول حاصل قابلیت کاربرد جهانی را بر روی بیشتر ماهواره‌ها و مناطق مختلف را نداشت. همچنین آنها در سال ۲۰۱۶ به بررسی تأثیر چهار نوع تصحیح جوی پرداختند که هیچکدام از این تصحیحات در حالت کلی رضایت بخش نبود. Kheireddine et al. (2018) به بررسی تأثیر جذب نور CDOM در لایه بالایی دریای قرمز پرداختند. آن‌ها از داده‌های فراطیفی در بازه ۴۰۰ تا ۷۴۰ نانومتر و الگوریتم رنگ اقیانوس استفاده کردند که در نتیجه تحقیقاتشان به این نتیجه رسیدند که استفاده از الگوریتم رنگ اقیانوس که با استفاده از روابط تجربی بیان شده است، میزان غلظت بالای CDOM را در نظر نمی‌گیرد و مناسب اندازه‌گیری Chla نمی‌باشد.

یکی دیگر از مشکلات موجود، با توجه به این که CDOM رفتار پیچیده و غیر خطی دارد، استفاده بیشتر محققان از روش رگرسیون خطی است که در بیشتر مواقع جواب مناسبی را نمی‌دهد (Kutser et al., 2015; Slonecker et al., 2005). یکی از روش مناسب برای حل این مشکل استفاده از الگوریتم‌های فراابتکاری است، اما به دلیل رفتار CDOM می‌تواند در الگوریتم‌های مختلف، جواب‌های متفاوتی به دست آورد. در این خصوص، Ruescas et al. (2018) با بکارگیری پنج روش یادگیری ماشین به مقایسه نتایج این پنج الگوریتم جهت

"واسکینا داچی"^۵ که شامل ۳۵۰ کیلومتر مربع از قسمت مرکزی شبه جزیره یامال (۷۰°N, ۷۰°E) در Yamalo-Nenets Autonomous Okrug واقع در شمال غربی سیبری است، استفاده می‌شود. لازم به ذکر است محدوده مطالعاتی در مجاورت یک ایستگاه تحقیقاتی با نام "واسکینا داچی" (۶۸°۵۴' N, ۷۰°۱۷' E) واقع گردیده است. محدوده وسیعی از این منطقه با دریای کارا، خلیج Baydaratskaya در غرب و با خلیج Ob در شرق هم مرز است. در بخش شمالی این شبه جزیره، تنگه Malygina و فراتر از آن جزیره بیلی قرار دارد. قسمت مرکزی منطقه مورد مطالعه با جزیره‌های کوچک و مرتفع مشخص می‌شود که به طور گسترده توسط دره‌های باریک رودخانه‌ای، جویبارهای کوچک و دره‌ها جدا می‌شوند.

۲-۲- داده‌های زمینی

داده‌های مکانی ضریب جذب CDOM دریاچه‌ها در طیف ۴۴۰ نانومتر در منطقه مطالعاتی حد فاصل سال‌های ۲۰۱۳ تا ۲۰۱۶ از سایت PANGAEA (<https://www.pangaea.de>) که ارائه‌کننده داده‌هایی برای علوم زمین و علوم زیست محیطی است، گرفته شده است. نمونه‌های آب در شرایط آب و هوایی آرام و توسط بطری نمونه‌گیری از ۳۰ سانتیمتری بالای آب، نزدیک به ساحل یا در مرکز دریاچه‌ها به وسیله قایق جمع‌آوری شده است.

بازیابی CDOM پرداختند. نتایج آنها نشان داد الگوریتم SVR و GPR در تصاویر ماهواره Sentinel 2,3 با ضریب تعیین بالایی مناسب بازیابی CDOM توسط این تصاویر می‌باشد.

با توجه به مطالبی ذکر شده، در این پژوهش از الگوریتم‌های SVR و SVM برای بازیابی و طبقه‌بندی CDOM از تصاویر ماهواره لندست ۸ استفاده شده است، همچنین تصحیح اتمسفری FLAASH تا حدی پاسخ مناسبی به بازیابی CDOM می‌دهد، پس از این تصحیح استفاده شده است (Kutser et al., 2016). ضمن اینکه تاکنون هیچ مقاله‌ای در مورد بازیابی CDOM با استفاده از الگوریتم GA-SVR گزارش نشده است. نتایج این پژوهش می‌تواند برای کنترل و حفاظت اکوسیستم دریاچه‌ها مورد استفاده قرار گیرد. شایان ذکر است که به علت عدم وجود داده‌های CDOM در ایران از داده‌های دریاچه‌های شمال سیبری استفاده شده است که این پژوهش می‌تواند سرآغازی برای توجه بیشتر محققان و مدیران به این موضوع در کشورمان ایران باشد.

۲- مواد و روش‌ها

۲-۱- منطقه مورد مطالعه

در این پژوهش از اطلاعات بیوشیمی دریاچه‌های دو منطقه "یامال"^۴

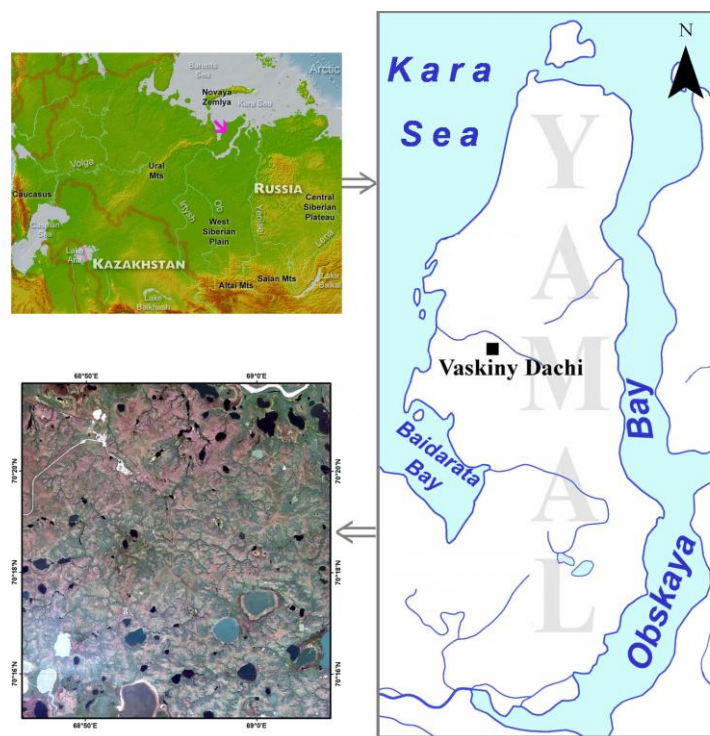


Fig. 1- Study area

شکل ۱- منطقه مورد مطالعه

از ارکان مهم تحقیقاتی که بر پایه طبقه‌بندی انجام می‌شوند، دقت در اجرای آن است؛ در این تحقیق از دقت کاپا و کلی استفاده می‌شود و همچنین تصاویر طبقه‌بندی شده به منظور سهولت در نمایش اطلاعات در بازه‌ی رنگی آبی تا قرمز نشان داده می‌شود که در این بازه رنگ آبی به کمترین میزان و رنگ قرمز به بیشتر میزان CDOM تعلق می‌گیرد (بدیهی است که مابقی مقادیر بین طیف رنگی این دو رنگ قرار خواهند گرفت). در نتیجه نقشه مکانی به دست آمده نمایانگر پراکندگی CDOM است.

۲-۵- رگرسیون بردار پشتیبان

الگوریتم SVR یکی از الگوریتم‌های یادگیری ماشین است که بر اساس تئوری آموزش آماری استوار است. این روش که جزء روش‌های آموزشی نظارت‌شده^۶ است که بین داده‌های ورودی و مقدار پارامتر وابسته، بر اساس کمینه‌سازی ریسک ساختاری ارتباط برقرار می‌کند. کمینه‌سازی ریسک ساختاری، نسبت به روش متداول کمینه‌سازی ریسک تجربی مورد استفاده در الگوریتم‌هایی مانند شبکه‌های عصبی و روش‌های کلاسیک آماری، دارای برتری است و برخلاف روش‌هایی همانند شبکه‌های عصبی، به جواب‌های موضعی همگرا نمی‌شود (Tu, 1996). SVR روشی برای تخمین یک تابع است که بر اساس داده‌های آموزشی از یک شیء ورودی به یک عدد واقعی نگاشت می‌شود. در مسائل رگرسیون، بردارهای ورودی به یک فضای چندبعدی، نگاشت می‌شوند؛ سپس یک ابرصفحه ساخته می‌شود که بردارهای ورودی را با بیشترین فاصله ممکن، از یکدیگر جدا می‌کنند. برای رفع مشکل انجام عملیات در فضایی با ابعاد زیاد، از یک تابع کرنل استفاده می‌شود. در این صورت عملیات می‌تواند با همان سرعت فضای داده‌های ورودی انجام شود. در واقع با استفاده از تابع کرنل، مشکل چندبعدی و غیرخطی بودن نگاشت، مرتفع می‌گردد. در حقیقت هدف SVR، تخمین پارامترهای وزن‌ها و شیب تابعی است که بهترین انطباق را بر داده‌ها داشته باشد (Yu and Kim et al., 2012). داده‌های ورودی، مقادیر بازتاب نسبت باندهای انتخابی و مقادیر میدانی ضریب جذب ماده آلی محلول در ۴۴۰ نانومتر می‌باشد.

نمونه‌های CDOM پس از نمونه‌گیری میدانی به‌طور مستقیم فیلتر شدند. فیلترهای CDOM با فیلتراسیون از طریق فیلترهای فیبری شیشه‌ای با اندازه ۰/۷ میکرومتر (واتمن) تهیه شدند و برای جلوگیری از تجزیه نوری CDOM، در شرایط سرد و تاریک نگهداری شدند (<https://www.pangaea.de>).

۲-۳- داده‌های سنجش از دور

تصاویر ماهواره لندست ۸ در سال‌های ۲۰۱۳ تا ۲۰۱۶ از سایت USGS اخذ گردید و سعی بر انتخاب تصاویر با کمترین پوشش ابری شده است. حداکثر فاصله زمانی بین اخذ تصاویر ماهواره‌ای و داده‌های زمینی برای دریاچه‌ها باید بین ۱ تا ۲ ماه باشد که این مهم در این پژوهش رعایت شده است (Brezonik et al., 2014). در ابتدا به تصحیحات مورد نیاز برای آماده‌سازی تصاویر ماهواره‌ای لندست ۸ به منظور بررسی امکان انجام پردازش‌های لازم پرداخته می‌شود. سازمان زمین‌شناسی آمریکا (USGS) تصاویر ماهواره لندست را به‌صورت زمین مرجع در سطح ۱ در اختیار کاربران قرار می‌دهد. تصحیحات رادیومتریک و اتمسفری بدلیل مقایسه داده‌های تصاویر ماهواره‌ای در زمان‌های متفاوت با استفاده از نرم‌افزار ENVI اصلاح گردیده‌اند (جدول ۱).

۲-۴- ساختار پژوهش

برای درک بهتر از روش انجام پژوهش و ساختار آن، فلوچارتی در شکل ۲ ارائه شده است. در مرحله پیش‌پردازش به تصحیحات مورد نیاز برای آماده‌سازی تصاویر ماهواره‌ای لندست ۸ جهت بررسی امکان انجام پردازش‌های آتی پرداخته می‌شود و پس از آن، الگوریتم‌های SVM (ماشین بردار پشتیبان) و SVR (رگرسیون بردار پشتیبان) جهت انجام پردازش استفاده می‌گردد. بدین ترتیب، ابتدا از الگوریتم بهینه شده SVR (GA-SVR) جهت انتخاب مناسب‌ترین نسبت باند به منظور تعیین ضریب جذب CDOM استفاده می‌شود. سپس نسبت باندی موردنظر بر روی تصاویر ماهواره‌ای پیاده‌سازی شده و با استفاده از الگوریتم SVM این تصاویر طبقه‌بندی می‌شوند. از آنجایی که یکی

Table 1- Specifications of the Landsat 8 images

جدول ۱- مشخصات تصاویر لندست ۸

Image ID	Satellite	Date	Path	Row
LC81650112013234LGN02	Landsat8	2013/07/22	165	011
LC81650112014237LGN01	Landsat8	2014/08/25	165	011
LC81670102015254LGN01	Landsat8	2015/09/25	167	010
LC81650122016227LGN01	Landsat8	2016/08/14	165	012

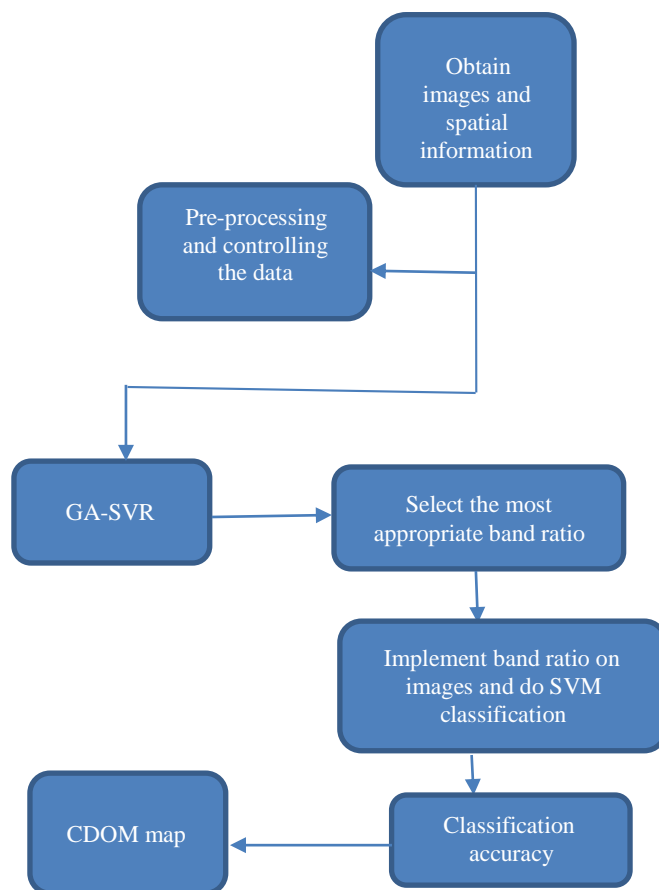


Fig. 2- Flowchart of the research
شکل ۲- نمودار فرآیند اجرای پژوهش

اعتبارسنجی شده است که k مساوی ۴ گرفته شده است و به جای استفاده از روش آزمون و خطا، پارامترهای SVR (C. σ. ε) با استفاده از الگوریتم ژنتیک بهینه‌سازی شده است (شکل ۳).

۲-۶- ماشین بردار پشتیبان

طبقه‌بندی SVM یک روش طبقه‌بندی باینری است که با تعیین یک صفحه تفکیک‌کننده بهینه در فضای ویژه داده‌های آموزشی، کلاس‌های مختلف را با حداکثر جدایی بین آن‌ها، تفکیک می‌نماید. فرآیندهای که حداکثر حاشیه را بین دو کلاس فراهم می‌نماید، فرآیندهای بهینه و داده‌های نزدیک‌تر به فرا صفحه، بردارهای پشتیبان نامیده می‌شوند (Vapnik, 1995).

ماشین بردار پشتیبان (SVM) یک روش یادگیری ماشین گسترده‌ای است که بر اساس اصل کمینه کردن ریسک ساختاری استفاده می‌شود و می‌تواند مشکل طبقه‌بندی نمونه‌های کوچک و غیرخطی را حل کند.

نسبت باندهای مورد استفاده در الگوریتم، نسبت‌بندی Coastal به قرمز (R_1/R_4)، قرمز به مادون قرمز نزدیک (R_4/R_5) و سبز به قرمز (R_3/R_4) می‌باشد. برای بهینه‌سازی پارامترهای الگوریتم SVR از الگوریتم ژنتیک استفاده شده است. نتایج به‌دست آمده از الگوریتم رگرسیون بردار پشتیبان به تعیین پارامترهای اولیه آن وابسته است که در این پژوهش سعی بر بهینه‌سازی مقادیر این پارامترها با استفاده از الگوریتم ژنتیک شده است. هدف از استفاده از این تابع مقایسه نسبت‌های باندهای موجود برای اندازه‌گیری ضریب جذب ماده آلی محلول رنگی در طیف ۴۴۰ نانومتر می‌باشد. هنگامی جستجو نتیجه بخش خواهد بود که به حداکثر نسل ممکن برسیم یا همگرایی حاصل شده باشد و یا معیارهای توقف برآورده شده باشد. مقدار جمعیت الگوریتم ژنتیک و تعداد تکرار را با توجه به نوع برازش و مقادیر خطا بدست آمده تنظیم شده است. به دلیل تعیین پارامتر برهم ریختگی داده‌ها در هر بار اجرا و اندازه‌گیری داده‌ها مقادیری مشخص و متفاوت به‌دست می‌آید، که در این رابطه سعی بر این شده که بهترین پارامترها در نظر گرفته شود. عملکرد الگوریتم از طریق اعتبارسنجی متقابل^۷

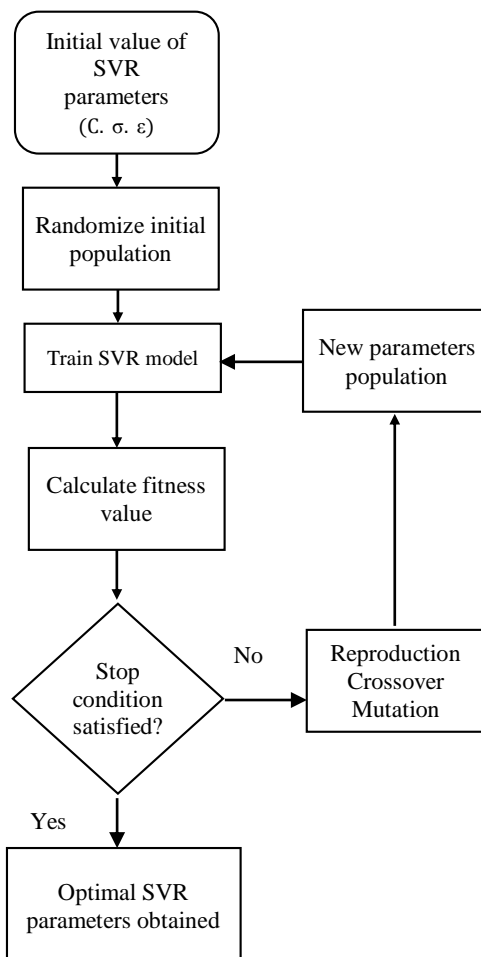


Fig. 3- Flowchart of the SVR algorithm
 شکل ۳- نمودار اجرای الگوریتم SVR

$$R^2 = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} \quad (1)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (2)$$

$$RMSE = \sqrt{MSE} \quad (3)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (4)$$

$$O.A. = \frac{\sum_{i=1}^k n_{ii}}{n} \quad (5)$$

$$K = \frac{n \sum_{i=1}^k n_{ii} - \sum_{i=1}^k n_{i+} n_{+i}}{n^2 - \sum_{i=1}^k n_{i+} n_{+i}} \quad (6)$$

اصل اساسی SVM این است که به دنبال یک ابرصفحه جداکننده بهینه بوده تا مسئله طبقه‌بندی از نظر خطی از هم جدا شوند (Yuan S-F et al., 2006, Lv Z et al., 2016). با استفاده از الگوریتم SVM به طبقه‌بندی تصاویر حاصل از پیاده‌سازی نسبت بانندی پرداخته می‌شود تا میزان تغییرات CDOM در دو سال متوالی مشخص گردد.

۷-۲- معیارهای ارزیابی

در این پژوهش برای ارزیابی دقت و کارایی الگوریتم‌ها، از ضریب تعیین (R^2)، ریشه میانگین مربعات خطا (RMSE)، میانگین مربعات خطا (MSE)، قدر مطلق خطا (MAE) در الگوریتم SVR استفاده شده است. ضریب تعیین به عنوان معیار اصلی در الگوریتم SVR در نظر گرفته شده و در طبقه‌بندی SVM از معیارهای کلی و کاپا برای تعیین دقت طبقه‌بندی استفاده گردیده است. روابط معیارهای ذکر شده به طبق روابط ۱ الی ۶ آورده شده است:

در رابطه‌های ۱ تا ۴، x, y نمایانگر مقادیر روی نمودار x, y, \hat{y}_i و Y_i به ترتیب مقادیر بدست آمده و مشاهداتی، n تعداد داده‌ها می‌باشند. در رابطه‌های ۵ و ۶ k تعداد کلاس‌ها، n تعداد کل پیکسل‌های معلوم در داده مرجع، n_{ij} اعضای قطری ماتریس ابهام، n_{i+} مجموع عناصر سطر i ام و n_{+i} مجموع عناصر ستون i ام هستند.

۳- نتایج و بحث

پیش‌بینی و اندازه‌گیری میزان تغییرات ماده آلی محلول رنگی در دریاچه‌ها اهمیت بخصوصی در سلامت اکوسیستم آبی دارد. به منظور دستیابی به این مهم، در این پژوهش از الگوریتم‌های ماشین بردار پشتیبان استفاده گردید. از این رو، سه نسبت باندی در طیف مرئی در طی سال‌های ۲۰۱۳ تا ۲۰۱۶ از تصاویر لندست ۸ در نظر گرفته شد و جهت ارزیابی این سه نسبت باندی از الگوریتم بهینه‌سازی شده رگرسیون بردار پشتیبان (GA-SVR) استفاده گردید. در الگوریتم رگرسیون از ۵۱ داده از دریاچه‌ها استفاده شده که ۸۰ درصد آن را برای آموزش و ۲۰ درصد برای آزمایش الگوریتم در نظر گرفته شده است و برای بهینه‌سازی پارامترهای الگوریتم ژنتیک، میزان جمعیت ۲۰ و تعداد تکرار ۳۰ انتخاب شده است. با توجه به مقادیر ورودی و پارامترهای تعیین شده، جدول ۲ که نشان دهنده نتایج ارزیابی مدل GA-SVR می‌باشد.

باتوجه به جدول ۲، داده‌های نسبت باندهای Coastal، سبز، قرمز و مادون قرمز نزدیک مورد بررسی قرار گرفته است. همانطور که در

جدول مشاهده می‌شود ضریب همبستگی ($R^2=0.71$) باندهای سبز به قرمز مقادیر مطلوب‌تری را نسبت به مقادیر نسبی دیگر باندها دارا می‌باشد. مقادیر RSME و MSE نسبت به سه باند دیگر کمتر می‌باشند ($RMSE=1/0.775 m^{-1}$, $MSE=1/1610 m^{-1}$). این نتایج هم برای داده‌های آموزش و هم برای داده‌های آزمایش در باند سبز و قرمز بهبود را نشان می‌دهد. هر چه طول موج افزایش پیدا می‌کند میزان جذب CDOM نیز بالاتر می‌رود. بنابراین، اثر CDOM بر بازتاب ماهواره لندست ۸ در باند Coastal قوی‌تر از باند سبز است. لذا دلایل متعددی وجود دارد که باند سبز برای بازیابی CDOM بهتر از باند آبی و Coastal است. اصلاح اتمسفری طول موج آبی به ویژه در مورد آب‌های کدر و یا دارای غلظت CDOM بالا همراه با مشکل است. در عین حال حساسیت سنسورهای آبی اغلب ناکافی است، به خصوص تابش خورشید در طول موج‌های آبی در مورد دریاچه‌های به رنگ قهوه‌ای (که غنی از CDOM می‌باشد) نزدیک به صفر است. جذب توسط فیتوپلانکتون نیز بر باند Coastal و آبی بیش‌تر از باند قرمز تأثیر می‌گذارد. بنابراین نسبت باند سبز/باند قرمز نسبت به نسبت باند Coastal/باند قرمز لندست ۸ برای بازیابی CDOM مناسب‌تر است. باند قرمز در اینجا برای اهداف نرمال کردن استفاده می‌شود؛ زیرا تأثیر CDOM در باند قرمز عملاً ناچیز است. در جدول ۳ مقادیر پارامترهای بهینه‌شده الگوریتم SVR با الگوریتم ژنتیک برای نسبت باندهای آورده شده است.

با توجه به داده‌های بدست آمده در منطقه Vaskiny Dachi در سال‌های ۲۰۱۳ تا ۲۰۱۶ در نسبت باندهای سبز به قرمز ماهواره

Table 2- Error values obtained from the ratio of Coastal, Green, Red, and Near Infrared bands

GA-SVR	Data Type	R ²	MSE	RMSE	MAE
R ₄ /R ₅	Train	0.14	3.5137	1.8745	1.6369
	Test	-0.10	2.5951	1.6109	1.2939
R ₁ /R ₄	Train	0.41	1.6380	1.2798	1.0346
	Test	-0.13	8.0103	2.8303	2.2933
R ₃ /R ₄	Train	0.81	0.8719	0.9338	0.5858
	Test	0.71	1.1610	1.0775	0.9464

Table 3- Optimized SVR model parameters for band ratios

SVR Model Parameters	C	ε	σ
R ₄ /R ₅	84.30	1.7696	0.2996
R ₁ /R ₄	1066.5	1.3904	0.3112
R ₃ /R ₄	524.6164	0.1099	0.6221

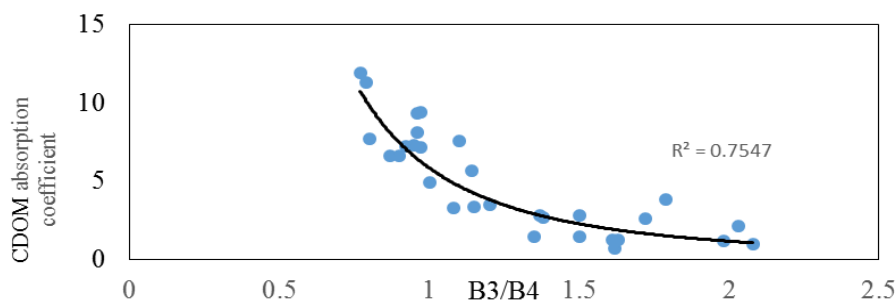


Fig. 4- A strong correlation between in situ α_{CDOM} and green/red band ratio at Vaskiny Dachi site in 2013-2016

شکل ۴- همبستگی قوی بین $\alpha_{CDOM}(440)$ میدانی و نسبت باند سبز به قرمز در منطقه Vaskiny Dachi در سال ۲۰۱۳ تا ۲۰۱۶

طبقه‌بندی شده از منطقه Vaskiny Dachi در سال‌های ۲۰۱۵ و ۲۰۱۴ (شکل ۵) با استفاده از طبقه‌بندی نظارت شده SVM جدول ۴ که نشان‌دهنده تعداد پیکسل‌های طبقه‌بندی شده در هر کلاس مشخص می‌باشد.

دقت کلی و ضریب کاپا این دو طبقه‌بندی به ترتیب ۰/۹۳ و ۰/۶ حاصل گردیده است. اختلافی که بیشتر در این طبقه‌بندی به چشم می‌خورد مربوط به اختلاف کلاس ۷ و ۸ می‌باشد و همانطور که در شکل ۵d دیده می‌شود نمایانگر افزایش مقدار ماده آلی محلول رنگی در دریاچه‌های شمال- شرقی تصویر در سال ۲۰۱۵ نسبت به سال ۲۰۱۴ می‌باشد. بیشتر پیکسل‌ها در کلاس ۱ تا ۳ قرار گرفته‌اند که با توجه به شکل ۵ به دلیل پوشش بیشتر تصویر توسط زمین منطقه می‌باشد.

لندست ۸ و داده‌های زمینی ضریب جذب CDOM در طیف ۴۴۰ نانومتر نمودار این دو متغیر نسبت به هم رسم شده است که محور x مقادیر نسبت باندهای سبز به قرمز ماهواره‌ی لندست ۸ می‌باشد و محور y نشان‌دهنده مقادیر ضریب جذب ماده آلی محلول رنگی در طیف ۴۴۰ نانومتر است (شکل ۴). با استفاده از نسبت باندهای ضریب جذب CDOM بدست آمده به بررسی میزان تغییرات $\alpha_{CDOM}(440)$ پرداخته می‌شود و تصویر $\alpha_{CDOM}(440)$ (نقشه مکانی ضریب جذب ماده آلی محلول رنگی) را می‌توان بدست آورد.

به منظور طبقه‌بندی تصاویر ماهواره لندست ۸ در این تحقیق از روش ماشین بردار پشتیبان (SVM) استفاده شده است که با توجه به شناخت از میانگین کلی ماده آلی محلول رنگی در منطقه Vaskiny Dachi در سال ۲۰۱۵، ۱۶ کلاس آموزشی معرفی شده است که در هر کلاس ۱۰ داده آموزشی موجود است. با توجه به دو تصویر لندست ۸

Table 4- Number of pixels classified in CDOM absorption coefficient images in Vaskiny Dachi region in 2014 and 2015 and their differences

جدول ۴- تعداد پیکسل‌های طبقه‌بندی شده در تصاویر ضریب جذب CDOM در منطقه Vaskiny Dachi در سال ۲۰۱۴ و ۲۰۱۵ و اختلافشان از یکدیگر

Year	2014	2015	Diff
Class 1	263464	83587	179877
Class 2	419655	239330	180325
Class 3	82354	202047	119693
Class 4	48551	161887	113336
Class 5	37631	80508	42877
Class 6	47817	51956	4139
Class 7	4095	10111	6016
Class 8	2642	4880	2338
Class 9	1321	2592	1271
Class 10	1417	1577	160
Class 11	264	770	506
Class 12	1824	891	933
Class 13	567	507	60
Class 14	81	253	172
Class 15	3234	1870	1364
Class 16	8153	6570	1583

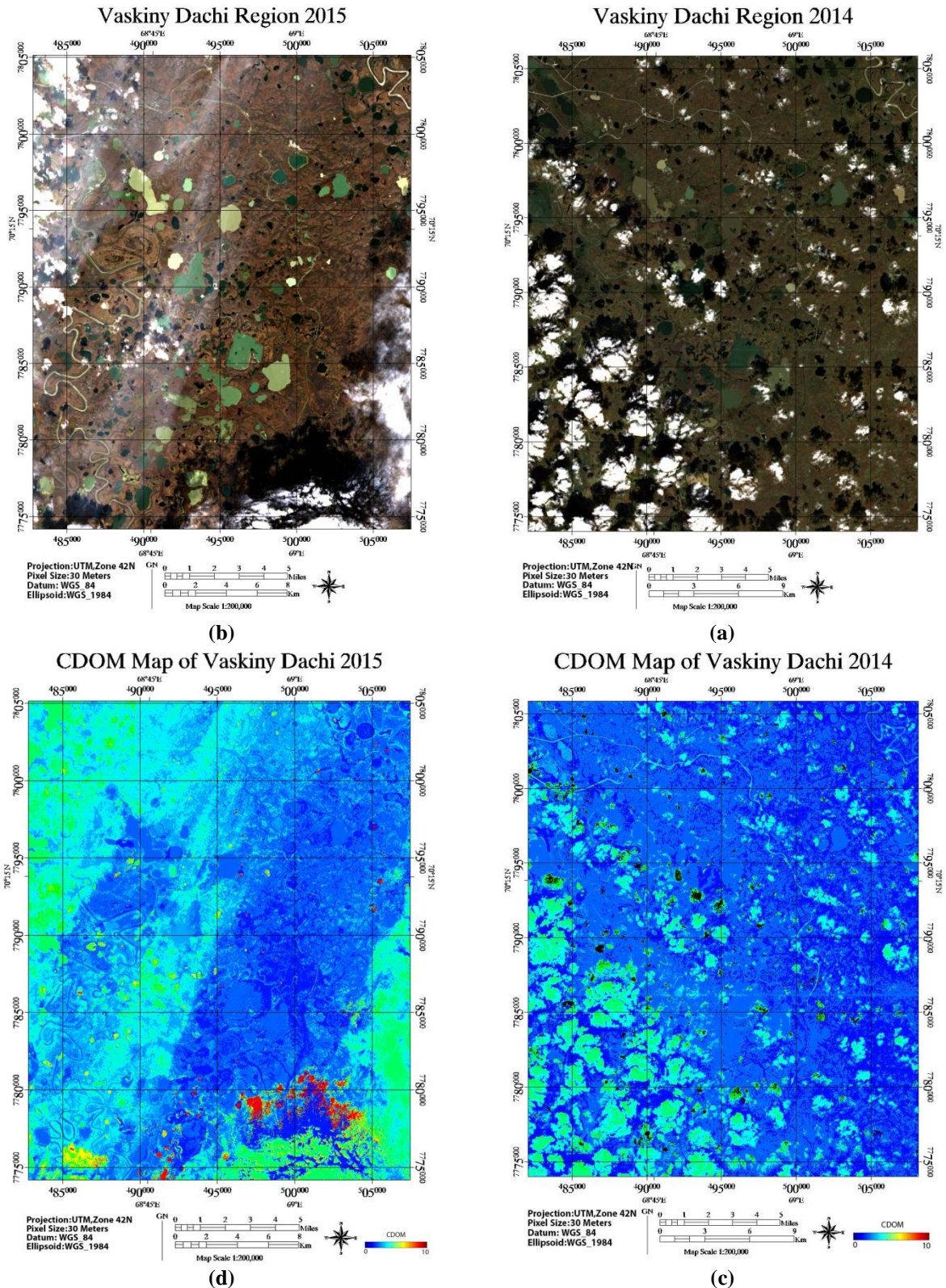


Fig. 5- The partial of Landsat 8 from the Vaskiny Dachi area in 2014 and 2015 (a, b) before classification, and (c, d) after classification

شکل ۵- تصویر قسمتی از لندست ۸ از منطقه Vaskiny Dachi در سال ۲۰۱۴ و ۲۰۱۵ (a,b) قبل از طبقه‌بندی و (c,d) بعد از طبقه‌بندی (تصاویر از بالا از چپ به راست به ترتیب a,b,c,d)

4m^{-1} تا 5m^{-1} می‌باشد. رودخانه‌های فرعی و کوچک که به صورت مویرگ هستند و در طیف آبی که پایین‌ترین مقدار α_{CDOM} در نظر گرفته شده است قرار دارند. میزان وضوح و تمایز این رودخانه‌ها در نقشه مکانی α_{CDOM} نمایان‌تر از تصویر بازتابی است. نقاط ریز قرمز که نشان‌دهنده میزان حداکثری در طیف تعیین شده می‌باشد، دریاچه‌های بسیار کوچکی در قسمت شمال شرقی منطقه و به دلیل تجمع بالا مواد آلی هیومیک هستند و در قسمت دریایی منطقه، تجمع آبی از نقاط قرمز دیده می‌شود که در نزدیکی ساحل قرار دارند، دلیل آن نیز ورود و تجمع مواد آلی در خط ساحلی توسط رودخانه‌ها است. در قسمت شرقی نقشه نقاط آبی بزرگتری وجود دارد که در تصویر بازتابی مناطق زراعی، به نظر می‌رسد که کمترین میزان طیف را به خود اختصاص داده‌اند. در خط ساحلی شرقی دریا میزان مقادیر CDOM افزایش یافته که می‌توان پراکندگی و تجمع ماده آلی محلول رنگی را به وضوح دید که این نمایش به علت وجود امواج دریایی و تجمع CDOM در خط ساحلی است. از این نقشه می‌توان برای تمایز مناطق آبی با تجمع کربن مختلف استفاده نمود.

در نقشه شکل ۶ تصویر دریای لاپتف در سال ۲۰۱۷ دیده می‌شود که برای بازه‌های مختلف α_{CDOM} (440) بازه‌های طیفی مختلفی اختصاص داده می‌شود، که این طیف رنگی انتخابی از آبی که کمترین میزان ضریب جذب CDOM را دارد تا قرمز که بیشترین مقدار را به خود اختصاص می‌دهد. این بدین مفهوم است که هر پیکسل که شامل مقدار α_{CDOM} (440) در تصویر لندست ۸ است به خود یک قسمتی از طیف رنگی بین آبی تا قرمز را به خود اختصاص می‌دهد. برای نمایش بهتر نقشه مکانی α_{CDOM} (440) طول بازه انتخابی مقادیر α_{CDOM} بین فاصله‌های 0m^{-1} تا 10m^{-1} انتخاب شده است. در این نقشه هدف نمایش بهتر میزان تجمع و پراکندگی α_{CDOM} در مناطق مختلف تصویر لندست ۸ می‌باشد.

در قسمت شمالی این نقشه مکانی، قسمتی بدون مقدار و همراه با تجمع آبی از نقاط قرمز (با کادر مشخص شده است) است که به دلیل تجمع ابری زیاد در آن قسمت است. در قسمت جنوب غربی تصویر رودخانه‌ای که به دریا می‌ریزد کاملاً مشخص است و با توجه طیف سبز رنگ آن می‌توان بیان کرد میزان ضریب جذب CDOM بین

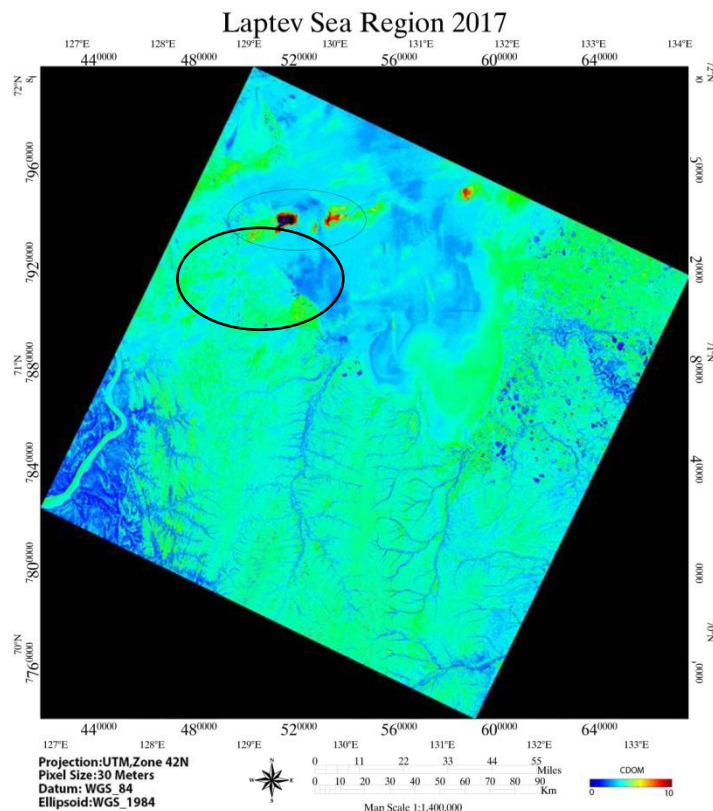


Fig. 6- Spatial map of colored dissolved organic matter absorption coefficient in the 440 nm spectrum in the Laptev Sea

شکل ۶- نقشه مکانی ضریب جذب ماده آلی محلول رنگی در طیف ۴۴۰ نانومتر در دریای لاپتف

۴- نتیجه گیری

اهمیت نقش دریاچه‌ها در اکوسیستم آبی و زیستگاه‌ها بر هیچ کس پوشیده نیست در نتیجه توجه بسیاری از محققان و اندیشمندان جهان بر این زمینه معطوف گردیده است. ضمن آنکه بر اساس بررسی‌ها و تحقیقات به عمل آمده، ماده آلی محلول (DOM) یکی از عوامل تاثیرگذار بر روی این اکوسیستم می‌باشد. این پژوهش تلاشی جهت ارائه اهمیت ماده آلی محلول رنگی (CDOM) در دریاچه‌ها و ارائه روشی بهینه و سریع برای اندازه‌گیری این ماده با استفاده از تصاویر ماهواره لندست ۸ بوده است. اجزای اصلی این پژوهش الگوریتم رگرسیون بردار پشتیبان (SVR) و الگوریتم ماشین بردار پشتیبان (SVM) بود که برای دستیابی به نتایج بهتر پارامترهای الگوریتم SVR توسط الگوریتم ژنتیک (GA-SVR) بهینه‌سازی شد. با توجه به مراحل اجرایی پژوهش، نتایج زیر به دست آمده است:

- نتایج حاصل از الگوریتم GA-SVR نمایانگر ضریب تعیین بالا و مقادیر کمینه خطای جذر میانگین مربعات، خطای میانگین مربعات، خطای مطلق در نسب باندی سبز/ قرمز می‌باشد ($RMSE=1/0.775 m^{-1}$, $MSE=1/161 m^{-1}$, $R^2=0/71$) و پس نسبت باندی سبز به قرمز مناسب‌ترین نسبت باندی در نظر گرفته شد که با توجه به تعیین پارامترهای خطا، این روش سرعت و دقت بهتری در این زمینه دارد.

- پس از پیاده‌سازی نسبت به دست آمده بر روی تصاویر لندست ۸، نقشه CDOM بدست آمد و سپس با استفاده از الگوریتم SVM، تصاویر حاصل به منظور تعیین میزان تغییرات این ماده در دو سال متوالی طبقه‌بندی گردید و به منظور شفاف‌سازی نتایج، بازه طیف رنگی مشخصی برای درک بهتر نقشه تعیین شد و میزان تغییرات بر اساس داده‌ها مشخص گردید. از این نتایج می‌توان جهت ارزیابی میزان و کنترل تغییرات CDOM در دریاچه‌ها استفاده کرد. لذا نویسندگان این پژوهش امیدوارند به دلیل اهمیت ماده آلی محلول رنگی (CDOM) در دریاچه‌ها و دیگر منابع آبی، این موضوع مورد توجه محققان دیگر نیز واقع گردد.

پی‌نوشت‌ها

- 1- Colored Dissolved Organic Matter
- 2- Dissolved Organic Matter
- 3- Phenolic
- 4- Yamal
- 5- Vaskiny Dachi
- 6- Supervised
- 7- K-Fold Cross Validation

۵- مراجع

- Babin M, Stramski D, Ferrari GM, Claustre H, Bricaud A, Obolensky G, and Hoepffner N (2003) Variations in the light absorption coefficients of phytoplankton, nonalgal particles, and dissolved organic matter in coastal waters around Europe. *Journal of Geophysical Research: Oceans*, Wiley Online Library 108(C7)
- Brezonik PL, Olmanson LG, Finlay JC, and Bauer ME (2015) Factors affecting the measurement of CDOM by remote sensing of optically complex inland waters. *Remote Sensing of Environment*, 157:199–215
- Budac D and Wan P (1992) Photodecarboxylation: Mechanism and synthetic utility. *Journal of Photochemistry and Photobiology A: Chemistry* 67(2):135–166
- Ferrari GM and Tassan S (1992) Evaluation of the influence of yellow substance absorption on the remote sensing of water quality in the Gulf of Naples: a case study. *International Journal of Remote Sensing* 13(12):2177–2189
- Häder D-P, Kumar HD, Smith RC, and Worrest RC (2007) Effects of solar UV radiation on aquatic ecosystems and interactions with climate change. *Photochemical & Photobiological Sciences*, Royal Society of Chemistry 6(3):267–285
- Jansson M (1998) Nutrient limitation and bacteria-Phytoplankton interactions in humic lakes. *Aquatic Humic Substances*, Springer 177–195
- Johannessen SC and Miller WL (2001) Quantum yield for the photochemical production of dissolved inorganic carbon in seawater. *Marine Chemistry* 76(4):271–283
- Karabashev GS, Khanaev SA, and Kuleshov AF (1993) On the variability of yellow substance in the Skagerrak and the Kattegat. *Oceanologica acta*. Gauthier-Villars 16(2):115–125
- Kheireddine M, Ouhssain M, Calleja ML, Morán XAG, Sarma YVB, Tiwari SP, and Jones BH (2018) Characterization of light absorption by Chromophoric Dissolved Organic Matter (CDOM) in the upper layer of the Red Sea. *Deep Sea Research Part I: Oceanographic Research Papers*, Elsevier 133:72–84
- Kutser T, Alikas K, Kothawala DN, and Köhler SJ (2015) Impact of iron associated to organic matter on remote sensing estimates of lake carbon content. *Remote Sensing of Environment* 156:109–116
- Kutser T, Pierson DC, Kallio KY, Reinart A, and Sobek S (2005) Mapping lake CDOM by satellite remote

- Dissolved Organic Matter (CDOM) retrieval with S2-MSI and S3-OLCI simulated data. *Remote Sensing, Multidisciplinary Digital Publishing Institute* 10(5):786
- Siegel DA, Maritorena S, Nelson NB, Hansell DA, and Lorenzi-Kayser M (2002) Global distribution and dynamics of colored dissolved and detrital organic materials. *Journal of Geophysical Research: Oceans* 107(C12):21
- Slonecker ET, Jones DK and Pellerin BA (2016) The new Landsat 8 potential for remote sensing of colored dissolved organic matter (CDOM). *Marine Pollution Bulletin* 107(2):518–527
- Tassan S (1988) The effect of dissolved “yellow substance” on the quantitative retrieval of chlorophyll and total suspended sediment concentrations from remote measurements of water colour. *Remote Sensing* 9(4):787–797
- Tranvik LJ (1992) Allochthonous dissolved organic matter as an energy source for pelagic bacteria and the concept of the microbial loop. *Dissolved organic matter in lacustrine ecosystems*, Springer, 107–114
- Tu J V (1996) Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology* 49(11):1225–1231
- Vähätalo A V, Salonen K, Salkinoja-Salonen M, and Hatakka A (1999) Photochemical mineralization of synthetic lignin in lake water indicates enhanced turnover of aromatic organic matter under solar radiation. *Biodegradation* 10(6):415–420
- Vapnik VN (1995) *The nature of statistical learning. Theory*, Springer
- Yu H and Kim S (2012) SVM tutorial-classification, regression and ranking. *Handbook of Natural computing* 1:479–506
- Yuan S-F and Chu F-L (2006) Support vector machines-based fault diagnosis for turbo-pump rotor. *Mechanical Systems and Signal Processing* 20(4):939–952
- sensing. *Remote Sensing of Environment* 94(4):535–540
- Kutser T, Tranvik L, and Pierson DC (2009) Variations in colored dissolved organic matter between boreal lakes studied by satellite remote sensing. *Journal of Applied Remote Sensing, International Society for Optics and Photonics* 3(1):33538
- Kutser T, Verpoorter C, Paavel B, and Tranvik LJ (2015) Estimating lake carbon fractions from remote sensing data. *Remote Sensing of Environment* 157:138–146
- Lv Z, Tang B, Zhou Y, and Zhou C (2016) A novel method for mechanical fault diagnosis based on variational mode decomposition and multikernel support vector machine. *Shock and Vibration, Hindawi* 2016
- Moss B (2012) Cogs in the endless machine: lakes, climate change and nutrient cycles: A review. *Science of the Total Environment, Elsevier* 434:130–142
- Nelson NB and Siegel DA (2013) The global distribution and dynamics of chromophoric dissolved organic matter. *Annual Review of Marine Science, Annual Reviews* 5:447–476
- Opsahl SP and Zepp RG (2001) Photochemically-induced alteration of stable carbon isotope ratios ($\delta^{13}C$) in terrigenous dissolved organic carbon. *Geophysical Research Letters, Wiley Online Library* 28(12):2417–2420
- Opsahl S and Benner R (1998) Photochemical reactivity of dissolved lignin in river and ocean waters. *Limnology and Oceanography* 43(6):1297–1304
- Organelli E, Bricaud A, Antoine D, and Matsuoka A (2014) Seasonal dynamics of light absorption by Chromophoric Dissolved Organic Matter (CDOM) in the NW Mediterranean Sea (BOUSSOLE site). *Deep Sea Research Part I: Oceanographic Research Papers, Elsevier* 91:72–85
- Ruescas AB, Hieronymi M, Mateo-Garcia G, Koponen S, Kallio K, and Camps-Valls G (2018) Machine learning regression approaches for Colored