

## کاربرد شبکه های عصبی مصنوعی (ANN) در تخمین محتوای آبی گیاهان (VWC) با استفاده از داده های فراطیفی

مژگان میرزایی<sup>۱\*</sup>، روشنگر درویش زاده<sup>۲</sup>، علی رضا شکیبا<sup>۳</sup>، علی اکبر متکان<sup>۴</sup>، متین شهری<sup>۵</sup>

<sup>۱</sup> کارشناس ارشد سنجش از دور و سیستم اطلاعات جغرافیایی - دانشگاه شهید بهشتی  
m\_mirzaie87@yahoo.com

<sup>۲</sup> استادیار گروه کارتوگرافی - دانشکده جغرافیا - دانشگاه تهران  
ravishh@alumni.itc.nl

<sup>۳</sup> استادیار دانشکده علوم زمین - گروه سنجش از دور و سیستم اطلاعات جغرافیایی - دانشگاه شهید بهشتی  
mypauk@yahoo.com

<sup>۴</sup> دانشیار دانشکده علوم زمین - گروه سنجش از دور و سیستم اطلاعات جغرافیایی - دانشگاه شهید بهشتی  
a-matkan@sbu.ac.ir

<sup>۵</sup> دانشجوی دکتری - دانشکده مهندسی عمران - دانشگاه علم و صنعت ایران  
matinshahri@yahoo.com

(تاریخ دریافت بهمن ۱۳۹۱، تاریخ تصویب آذر ۱۳۹۳)

### چکیده

با گسترش سنجش از دور فراطیفی امکان بهره گیری از گروه جدیدی از شاخص های طیفی و مدل های آماری، برای تخمین پارامترهای بیو فیزیکی و بیوشیمیایی گیاهان به وجود آمده است. یکی از پارامتر های بیوشیمیایی گیاه، محتوای آبی گیاه (VWC) است که پارامتری مهم در بخش کشاورزی است و می تواند در جهت آبیاری صحیح و ارزیابی شرایط خشکسالی مورد استفاده قرار بگیرد. در این تحقیق با استفاده از شبکه های عصبی مصنوعی و داده های فراطیفی که با استفاده از دستگاه اسپکترومتر GER 3700 بدست آمده اند به تخمین محتوای آبی گیاهان پرداخته شد. چهار گونه گیاهی متفاوت برای نمونه برداری طیفی و اندازه گیری محتوای آبی گیاه انتخاب شدند. در مجموع با ایجاد تغییر در تراکم تاج پوشش گیاهان ۹۵ نمونه شامل بازتاب های طیفی و همچنین اندازه گیری های محتوای آبی گیاهان برای آموزش شبکه مهیا شد. سپس یک شبکه پس انتشار خطا (پرسپترون چند لایه) با سه گروه از ورودی ها که شامل تمامی باند های باریک موجود، ده مولفه اصلی اول و همچنین چهار شاخص باریک باند گیاهی بودند برای تخمین VWC مورد استفاده قرار گرفت. برای ارزیابی کارایی شبکه در تخمین VWC از روش Cross-validation استفاده شد. با استفاده از رگرسیون خطی ساده بین خروجی شبکه و داده های اندازه گیری شده VWC گیاه و بر حسب معیارهای  $R_{cv}^2$  و  $RMSE_{cv}$  بهترین مدل انتخاب شد. نتایج تحقیق نشان دادند که شبکه های عصبی دارای قابلیت بالا در تخمین VWC با استفاده از داده ای فراطیفی می باشند ( $R_{cv}=0.88$ ,  $RMSE_{cv}=0.31$ ).

**واژگان کلیدی:** فراطیفی، شبکه عصبی مصنوعی، پس انتشار خطا، VWC

\* نویسنده رابط

## ۱- مقدمه

محتوای آبی گیاه (VWC)<sup>۱</sup> به عنوان شاخصی مهم در کاربرد های مختلف کشاورزی و جنگل داری مطرح است [۱۷]. بسیاری از فرآیندهای بیوشیمیایی در گیاهان از جمله فتوسنتز، تبخیر و تعرق و تولید اولیه مستقیماً با میزان آب گیاه در ارتباط است [۱۱]. VWC علاوه بر کاربرد های گسترده ای که در بخش کشاورزی، جنگل داری و هیدرولوژی دارد، می تواند برای ارزیابی رطوبت خاک نیز مورد استفاده قرار بگیرد [۲۰]. روش های مبتنی بر سنجش از دور به عنوان تکنیک هایی قابل اطمینان، سریع و غیر مخرب برای تخمین متغیر های فیزیکی و بیوشیمیایی گیاهان شناخته شده اند [۳،۴].

اغلب تولیدات سنجش از دور برای ارزیابی و تخمین پارامتر های بیوفیزیکی و بیوشیمیایی گیاهان از سنجنده هایی با باندهای پهن همانند NOAA<sup>۲</sup>، AVHRR<sup>۳</sup>، SPOT<sup>۴</sup> و TM/ETM<sup>۵</sup> سنجنده لندست استفاده می کنند که اغلب از سه الی هفت باند تشکیل شده اند [۹]. در چند دهه گذشته ابزار های سنجش از دور گیاهی بطور چشم گیری پیشرفت داشته اند. با گسترش سنجش دور فراطیفی امکان بهره گیری از گروه جدیدی از مدل های چند و تک متغیره، برای تخمین پارامترهای بیوفیزیکی و بیوشیمیایی گیاهان به وجود آمده است. از محدودیت های اساسی در تولیدات سنجش از دور پهن باند استفاده از میانگین اطلاعات طیفی در پهنای یک باند عرضی می باشد. استفاده از میانگین بازتابها در یک باند عرضی موجب از دست رفتن اطلاعات مهم و موثر که در باندهای باریک تر موجود می باشند، می شود [۹]. توسعه در زمینه سنجش از دور فراطیفی، قابلیت هایی برای غلبه بر این مشکل فراهم آورده است. تصویر برداری اسپکترومتری یا سنجش از دور فرا طیفی با سنجنده هایی متشکل از چند صد باند باریک و پیوسته طیفی در سرتاسر باندهای طیفی مابین ۴۰۰ تا ۲۵۰۰ نانومتر، دارای پتانسیل بالایی

در اندازه گیری پارامتر های گیاهی نسبت به استفاده از سنجنده هایی چند طیفی می باشند [۱۲].

روش های سنجش از دوری برای تخمین پارامتر های گیاهی به طور کلی به دو گروه: تکنیک های آماری (روش های استدلالی<sup>۶</sup> و یا روش های تجربی<sup>۷</sup>) و روش های فیزیکی (یا روش های استنتاجی<sup>۸</sup>) طبقه بندی می شوند [۱۲]. برای تخمین مقدار آب گیاهان تا کنون روش های گوناگونی از جمله روش های آماری [۸،۱۷] و روش های فیزیکی [۵،۶،۱۰،۱۱،۲۴،۲۸] مورد استفاده قرار گرفته است. تکنیک های آماری برای بدست آوردن همبستگی بین پارامتر گیاهی و بازتاب طیفی آن و یا شاخص های گیاهی مورد استفاده قرار می گیرند [۱۲].

روش های آماری از پرکاربرد ترین روش های مورد استفاده در تخمین پارامتر های گیاهی در سنجش از دور می باشد. این روش ها شامل: روش های تک متغیره<sup>۹</sup> (محاسبه شاخص های طیفی گیاهی) و مدل های چند متغیره<sup>۱۰</sup> (از جمله رگرسیون خطی مرحله ای (SMR)<sup>۱۱</sup>، رگرسیون حداقل مربعات جزئی (PLSR)<sup>۱۲</sup>، شبکه های عصبی مصنوعی (ANN)<sup>۱۳</sup> و غیره می باشند. روش های تک متغیره اغلب شامل شاخص های گیاهی می باشند. شبکه های عصبی مصنوعی از جمله روش های چند متغیره آماری است که می تواند از تمامی اطلاعات موجود در باند ها استفاده کند. از مزایای شبکه های عصبی به ویژگی غیر پارامتریک بودن آن می توان اشاره کرد به این معنی که توزیع داده ها در شبکه مهم نمی باشد. توانایی شبکه های عصبی در ارتباط برقرار کردن مابین اطلاعات پیچیده طیفی و ویژگی های گیاهی مورد نظر آنها را به عنوان ابزاری ایده آل برای توصیف روابط پیچیده و غیر خطی که مابین پارامتر های گیاهان و خصوصیات طیفی در سطح تاج پوشش موجود می باشند، تبدیل ساخته است [۲۶]. اگر چه در مطالعات اولیه شبکه های عصبی اغلب برای

۶ Inductive

۷ Empirical

۸ Deductive

۹ Univariate technique

۱۰ Multivariate technique

۱۱ Stepwise Multiple Linear Regression

۱۲ Partial Least Square Regression

۱۳ Artificial Neural Network

۱ Vegetation Water Content

۲ National Oceanic and Atmospheric Administration

۳ Advanced Very High Resolution Radiometer

۴ Satellite Pour l'Observation de la Terre

۵ Thematic Mapper/ Enhanced Thematic Mapper

روشن و تیره) داده شد. نهایتاً ۹۵ نمونه آماده شد، ۲۴ نمونه برای هر گونه (یک نمونه حذف شد). ۴۸ نمونه در خاک تیره و ۴۷ نمونه در خاک روشن. بدلیل وجود نویز بالا محدوده های طیفی کمتر از ۴۰۰ نانو متر و بیشتر از ۲۴۰۰ نانومتر حذف شدند. فیلتر Savitzky-Golay [۲۳]، با اندازه پنجره ۱۷ پیکسل و چند جمله ای درجه دو، برای از بین بردن نویز های احتمالی دستگاه اندازه گیری طیفی بر روی ماتریس بازتاب های طیفی اجرا شد. پردازش ها و تحلیل های انجام شده تماماً با استفاده از نرم افزار MATLAB 7.8 انجام گرفتند.

## ۲-۲- شبکه عصبی مصنوعی

یک شبکه عصبی مصنوعی یک ماشین محاسباتی است که از ساختار سیستم عصبی بیولوژیکی الهام گرفته اند. پایه ای ترین قسمت ساختمان یک شبکه عصبی نرون ها هستند. نرون یک واحد پردازش است که ورودی را از منبع دریافت کرده و با تعریف رفتاری غیر خطی در جهت رسیدن به شبکه، به خروجی نهایی می رسد [۲۲].

الگوریتم آموزشی پس انتشار یکی از پرکاربرد ترین انواع شبکه های عصبی است که در این تحقیق نیز مورد استفاده قرار گرفته است. هدف اصلی یک شبکه عصبی ساخت مدلی است با استفاده از ورودی ها، به طوری که این شبکه آموزش دیده، قابلیت پیش بینی خروجی ها با استفاده از ورودی های جدید که شبکه تاکنون با آنها مواجه نشده است را داشته باشد. الگوریتم پس انتشار یک شبکه عصبی چند لایه است که متشکل از یک لایه ورودی، چند لایه پنهان، و یک لایه خروجی می باشد. آموزش یک شبکه عصبی توسط تطبیق وزن ها به گونه ای که خطای شبکه که اختلاف بین خروجی شبکه و مقادیر واقعی است انجام می شود. در انتهای فاز آموزش، شبکه مدلی ارائه می کند که قادر به پیش بینی مقادیر خروجی توسط ورودی ها باشد.

پایه ای ترین قسمت ساختمان یک شبکه عصبی نرون ها هستند. نرون یک واحد پردازش که دارای چند ورودی (معمولاً بیشتر از ۱) و یک خروجی است. نرون ها می توانند تک ورودی یا چند ورودی باشند. معمولاً یک نرون بیش از یک ورودی دارد.

طیفه بندی مورد استفاده قرار می گرفتند، اما این تکنیک همچنین پتانسیل بالایی در پیش بینی متغیر ها نیز نشان داد [۲۶].

شبکه های عصبی مصنوعی برای تخمین پارامتر های گیاهی با استفاده از سنجش از دور در مطالعات پیشین برای تخمین شاخص سطح برگ [۲]، ارزیابی میزان کلروفیل محصول [۱۹]، تخمین میزان محصول [۱۶،۲۷]، تخمین رطوبت خاک [۷،۱۴،۲۱] و تخمین بیوماس گیاهان [۱۸]، مورد استفاده قرار گرفته است.

لذا این تحقیق در نظر دارد تا توانایی شبکه های عصبی را در برآورد محتوای آبی گیاهان در سطح تاج پوشش گیاه را با استفاده از داده های فراطیفی مورد ارزیابی قرار دهد و همچنین با استفاده از مولفه های اصلی و همچنین شاخص های گیاهی باریک باند به عنوان ورودی شبکه از ایجاد مشکل بیش برآزش جلوگیری کند.

## ۲- مواد و روش ها

### ۲-۱- اندازه گیری های آزمایشگاهی و پیش پردازش داده ها

به منظور نمونه برداری چهار گونه گیاهی انتخاب شدند. گونه های گیاهی در شکل و اندازه های برگ متفاوت بودند. در مجموع ۲۴ گیاه (از هر گونه ۶ نمونه) در شرایط گلخانه ای جمع آوری گردید [۱۳]. اندازه گیری های طیفی تاج پوشش گیاهان در آزمایشگاه کنترل شده سنجش از دور انجام گرفت. در محل اندازه گیری دیوار ها و سقف با مواد سیاه رنگ مات به منظور جلوگیری از بازتاب های زائد، پوشیده شدند. اندازه گیری ها با استفاده از دستگاه اسپکترومتر GER 3700 صورت گرفت. محدوده طول موجهای این اسپکترومتر بین ۳۵۰ تا ۲۵۰۰ نانومتر می باشد که در محدوده طیفی ۱۰۵۰-۳۵۰۰ فاصله طول موج ها ۱،۵ نانومتر، در فاصله ۱۹۰۰-۱۰۵۰ دارای فاصله طیفی ۶،۲ نانو متر و در ناحیه ۲۵۰۰-۱۹۰۰ درای فواصل ۹،۵ نانومتری هستند. برای ایجاد تنوع در اندازه گیری های طیفی و تولید نمونه های بیشتر که حداکثر اختلاف را دارا باشند، تغییراتی در حجم تاج پوشش گیاه و همچنین خاک پس زمینه (متشکل از خاک

## ۲-۲-۱- الگوریتم پس انتشار خطا

محاسبه می‌شود. سپس این خطا در جهت مخالف حالت پیش رونده داخل شبکه انتشار می‌یابد. در این مرحله وزن‌ها در جهتی که به مقادیر واقعی نزدیک شوند تنظیم می‌شوند. در مرحله بازگشتی، وزن‌های نرون‌ها با شروع از لایه خارجی به سمت لایه ورودی شروع به تنظیم شدن می‌کنند که به صورت حرکت به سمت چپ از لایه آخر به سمت لایه اول می‌باشد [۲۵]. هدف از تنظیم وزن‌ها کاهش خطای شبکه است که معمولاً خطای حداقل میانگین مربعات<sup>۵</sup> معیار خطا می‌باشد و بصورت زیر تعریف می‌شود:

$$E(w) = 1/2 \sum_{j,k} (a_{j,k} - o_{j,k})^2 \quad (۳)$$

که  $w$  وزن،  $a_{j,k}$  مقدار خروجی نرون  $j$ ام در لایه آخر که از نمونه آموزشی  $k$ ام بدست آمده و  $o_{j,k}$  مقدار واقعی نرون  $j$ ام در نمونه آموزشی  $k$ ام می‌باشد. برای حداقل سازی مقدار  $E(w)$  روش‌های متفاوتی وجود دارد که یکی از پرکاربردترین روش‌ها روش سریع‌ترین کاهش شیب<sup>۶</sup> می‌باشد که در این روش وزن‌ها با استفاده از رابطه زیر تنظیم می‌شوند:

$$w_{ji}^{n+1} = w_{ji}^n - \eta \frac{\partial E}{\partial w_{ji}^n} \quad (۴)$$

به طوری که  $w_{ji}^n$  وزن نرون  $i$ ام به نرون  $j$ ام در زمان  $n$  است،  $\eta$  نرخ آموزش بوده که مقداری است ثابت بین ۰ و ۱ و  $\frac{\partial E}{\partial w_{ji}^n}$  شیب (مشتق) در نقطه  $w_{ji}^n$  می‌باشد که می‌تواند به عنوان تغییر مقدار خطای  $E$  که ناشی از تغییر وزن  $w_{ji}^n$  می‌باشد در نظر گرفته شود.

مقادیر اولیه وزن‌ها معمولاً مقادیری کوچک و تصادفی هستند. مقدار شیب در معادله فوق بصورت زیر بسط داده می‌شود:

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial a_i} \frac{\partial a_i}{\partial x_j} \frac{\partial x_j}{\partial w_{ji}} = \frac{\partial E}{\partial a_i} f'(x_j) a_i = \delta_j a_i \quad (۵)$$

که  $f'(x_j)$  مشتق اول تابع انتقال که معمولاً تابعی است سیگموئیدی،  $a_i$  مقدار ورودی نرون و  $\delta_j$  مقدار

الگوریتم پس انتشار تعمیم قانون یادگیری Widrow-Hoff برای شبکه‌های چند لایه و توابع انتقال جدایی پذیر می‌باشد این روش از روشهای نظارت شده است و از "قانون دلتا"<sup>۱</sup> یا "الگوریتم کاهش شیب خطا"<sup>۲</sup> برای کم کردن خطا استفاده می‌کند. که در این الگوریتم وزن‌ها در خلاف جهت شیب در تابع خطا شروع به اصلاح شدن می‌کند. واژه پس انتشار به رفتاری برمی‌گردد که در آن شیب برای شبکه‌های چند لایه غیر خطی محاسبه می‌شود [۱۵].

الگوریتم پس انتشار به دو مرحله: مرحله پیش‌رونده<sup>۳</sup> و مرحله بازگشتی<sup>۴</sup> تقسیم می‌شود. در مرحله پیش‌رونده، هر یک از نمونه‌ها به شبکه وارد شده و مقادیر نرون‌ها بطور متوالی از لایه ورودی به سمت لایه خروجی به هنگام می‌شود. در مرحله پیش‌رونده، هر یک از نمونه‌ها به شبکه وارد شده و مقادیر نرون‌ها بطور متوالی از لایه ورودی به سمت لایه خروجی به هنگام می‌شود. در این مرحله تمام مقادیر ورودی به یک نرون بصورت زیر محاسبه می‌شود:

$$X_j = \sum_i a_i w_{ji} \quad (۱)$$

که  $a_i$  مقدار نرون  $i$ ام و  $w_{ji}$  وزن اتصال نرون  $i$ ام به نرون  $j$ ام می‌باشد. بعد از محاسبه با استفاده از فرمول فوق با اعمال تابع فعال سازی  $f$  مقدار  $x_j$  به مقدار خروجی آن نرون (ورودی نرون لایه بعدی) تبدیل می‌شود.

$$f(x_j) = a_j \quad (۲)$$

تابع فعال‌سازی تانژانت-سیگموئیدی اغلب در این شبکه‌ها مورد استفاده قرار می‌گیرد. این شبکه‌ها همچنین شبکه‌های پرسپترون چندلایه نیز نامیده می‌شود. این شبکه‌ها با استفاده از الگوریتم پس انتشار خطا آموزش دیده می‌شود. در مرحله پیش‌رونده وزن‌ها تنظیم نمی‌شوند و ثابت می‌مانند. هنگامی که خروجی بدست آمدند با مقادیر واقعی مقایسه می‌شوند و مقدار خطا

<sup>۱</sup> Delta Rule

<sup>۲</sup> Gradient Descent Algorithm

<sup>۳</sup> Forward

<sup>۴</sup> Backward

<sup>۵</sup> Least mean square error

<sup>۶</sup> Steepest descent

برای جلوگیری از احتمال ایجاد بیش برآزش در شبکه دو راه حل انتخاب شد:

- یکی تقسیم داده‌ها به نمونه‌های آموزشی، تست و اعتبارسنج.
- روش دیگر کاهش تعداد متغیرها یا باندهای طیفی بود. به این معنی که می‌توان از روش‌های فشرده سازی داده همانند مولفه‌های اصلی و همچنین استفاده از شاخص های طیفی برای ورود به شبکه به جای استفاده از تمامی باندها استفاده شد.

از این رو برای انتخاب بهترین شبکه و بهترین الگوریتم آموزش که قادر به تخمین VWC که قادر به میزان کمترین میزان خطا باشد ۳ گروه وردی شامل تمامی باند های طیفی موجود، ۱۰ مولفه اصلی اول و ۴ شاخص باریک باند طیفی بهینه شامل NDWI، SAVI2 و TSAVI به عنوان وردی شبکه مورد استفاده قرار گرفت که در مطالعه پیشین میرزایی و همکاران در سال ۱۳۹۰ محاسبه شده بودند. با الگوریتم‌های مختلف آموزش به شبکه ارائه شدند. تابع Tansig برای لایه میانی و تابع linear برای تابع خروجی انتخاب شد. این دو تابع به دلیل پیوسته و مشتق پذیر بودن در الگوریتم پس انتشار بسیار مورد استفاده قرار می‌گیرند [۱۵]. در این تحقیق از روش Cross validation برای ارزیابی دقت مدل‌ها استفاده شد. استفاده از این روش در شبکه‌های عصبی بدین شکل بود که هر شبکه با ۹۴ نمونه آموزش داده شد. سپس نمونه باقی مانده که در آموزش شرکت نداشت با استفاده از این شبکه شبیه‌سازی شد. پس از تکرار این مراحل به تعداد ۹۵ مرتبه، تعداد ۹۵ نمونه شبیه‌سازی شده بدست آمد. سپس نمونه‌های شبیه سازی شده که در واقع مقادیر تخمین زده شده VWC بودند در مقابل مقادیر اندازه‌گیری شده VWC قرار گرفته و مقدار  $RMSE_{CV}$  و  $R^2_{CV}$  شبکه محاسبه شد.

### ۳ - نتایج و بحث

#### ۳-۱- استفاده از شبکه عصبی و تمامی باند های طیفی به عنوان ورودی برای برآورد VWC

خطای نرون زام می باشد. مقدار خطا برای نرون های لایه آخر باشد معادله فوق بصورت زیر ساده می شود:

$$\delta_j = (a_j - o_j)f'(x_j) \quad (6)$$

به طوری که  $O_j$  مقدار واقعی و  $O_j$  مقدار پیش بینی شده می باشد.

از ترکیب معادلات و وزن ها برای لایه خروجی تنظیم می شود. برای سایر لایه ها مقدار  $\delta_j$  با استفاده از معادله زیر بدست می آید:

$$\delta_j = \left( \sum_K \delta_k w_{kj} \right) f'(x_j) \quad (7)$$

که  $k$  نشان دهنده نرون هایی هستند که مقادیر خروجی را از لایه زام در یافت می کند.

این قانون برای تنظیم وزن ها همراه با اتصالات شبکه به عنوان قانون تعمیم یافته دلتا<sup>۱</sup> نیز شناخته می شود [۲۵]. برخی مواقع پارامتر دیگری با نام مومنتوم نیز به معادله اضافه می شود :

$$w_{ji}^{n+1} = w_{ji}^n - \eta \frac{\partial E}{\partial w_{ji}^n} + \xi (w_{ji}^n - w_{ji}^{n-1}) \quad (8)$$

مقدار مومنتوم بین ۰ و ۱ بوده و برای جلوگیری از نوسان در یافتن مقدار مینیمم در سطح خطا به کار برده می شود که منجر به افزایش سرعت همگرایی می شود. شبکه های پی انتشار با چندین فاکتور کنترل می شوند: پارامتر های همراه با مدل، توپولوژی شبکه و نمونه های آموزشی. انتخاب یک شبکه بهینه اغلب کار دشواری می باشد [۲۵].

در این تحقیق یک شبکه پس انتشار خطا (پرسپترون چند لایه) با یک لایه ورودی، یک لایه پنهان و یک لایه خروجی مناسب تشخیص داده شد. تعداد نرون های ورودی به تعداد مولفه های ورودی ها بود. تعداد نرون های لایه پنهان مناسب برای هر یک از مجموعه ورودی ها انتخاب شدند. و نرون های لایه خروجی نیز تنها یک نرون بود که همان مقدار VWC به عنوان خروجی بود. از آنجایی که ۹۵ نمونه اندازه‌گیری شده VWC برای گیاهان موجود بود لذا از آموزش نظارت شده استفاده شد.

<sup>۱</sup> Generalized delta rule

این شبکه با نرون های متفاوت شبکه با ۲۵ نرون در لایه پنهان بهتر از سایر شبکه ها آموزش داده شد ( $R^2_{cv}=0.84$ )  $RRMSE_{cv}=0.36$ ،

الگوریتم های آموزشی متفاوتی برای شبکه هنگامی که تمامی باندهای طیفی (۵۸۴ باند طیفی) به عنوان ورودی در نظر گرفته بودند مورد بررسی قرار گرفتند که از میان آنها شبکه تنها با الگوریتم Scg به خوبی آموزش دید.

جدول ۱- شبکه های عصبی مختلف و نتایج بدست آمده از آنها در تخمین VWC با استفاده از تمامی باند های طیفی به عنوان ورودی

| ورودی ها               | شبکه بهینه | تابع فعال سازی  | تابع آموزش | $R^2_{cv}$ | $RRMSE_{cv}$ |
|------------------------|------------|-----------------|------------|------------|--------------|
| تمام باندها (۵۸۴ باند) | ۵۸۴*۱۵*۱   | {tansig,linear} | Scg        | ۰,۶۹       | ۰,۵۰         |
|                        | ۵۸۴*۱۷*۱   | {tansig,linear} | Scg        | ۰,۷۹       | ۰,۴۲         |
|                        | ۵۸۴*۲۰*۱   | {tansig,linear} | Scg        | ۰,۸۰       | ۰,۴۰         |
|                        | ۵۸۴*۲۵*۱   | {tansig,linear} | Scg        | ۰,۸۴       | ۰,۳۶         |
|                        | ۵۸۴*۳۰*۱   | {tansig,linear} | Scg        | ۰,۷۵       | ۰,۴۶         |

مولفه های اصلی بودند، مورد استفاده و مقایسه قرار گرفت. همانطور که در جدول (۲) نیز مشخص است، شبکه آموزش دیده شده با استفاده از الگوریتم آموزشی شیب توام (scg) در مقایسه با سایر الگوریتم های آموزشی مورد استفاده، دارای  $R^2$  بالاتر و RMSE پایین تر می باشد ( $R^2_{cv}=0.88, RRMSE_{cv}=0.31$ ). در واقع با این الگوریتم شبکه بهتر آموزش دیده.

### ۳-۲- استفاده از شبکه عصبی و ۱۰ مولفه اصلی به عنوان ورودی برای برآورد VWC

برای اطمینان از عدم بیش برآزش شبکه از مولفه های اصلی نیز به عنوان ورودی در شبکه عصبی استفاده شد. ساختار شبکه شامل یک لایه ورودی با ۱۰ نرون، یک لایه مخفی با ۶ و ۱۵ نرون و همچنین یک لایه خروجی با یک نرون استفاده شد. الگوریتم های آموزشی پس انتشار ارتجاعی (Trainrp)، الگوریتم شیب توام (Trainscg)، الگوریتم شبه نیوتن (Trainoss) و الگوریتم Levenbery-Marqwardt برای آموزش شبکه، هنگامی که ورودی

جدول ۲- شبکه های عصبی مختلف و نتایج بدست آمده از آنها در تخمین VWC با استفاده مولفه های اصلی به عنوان ورودی

| ورودی ها            | شبکه بهینه | تابع فعال سازی  | تابع آموزش | $R^2_{cv}$ | $RRMSE_{cv}$ |
|---------------------|------------|-----------------|------------|------------|--------------|
| ۱۰ مولفه اصلی (PCs) | ۱۰*۶*۱     | {tansig,linear} | Oss        | ۰,۷۰       | ۰,۴۹         |
|                     | ۱۰*۶*۱     | {tansig,linear} | Rp         | ۰,۸۶       | ۰,۳۳         |
|                     | ۱۰*۱۵*۱    | {tansig,linear} | Rp         | ۰,۵۹       | ۰,۶۴         |
|                     | ۱۰*۶*۱     | {tansig,linear} | Lm         | ۰,۶۵       | ۰,۵۴         |
|                     | ۱۰*۶*۱     | {tansig,linear} | Scg        | ۰,۸۸       | ۰,۳۱         |

### ۳-۳- استفاده از شبکه عصبی و چهار شاخص باریک باند به عنوان ورودی برای برآورد VWC

در تحقیق پیشین میرزایی و همکاران در سال ۱۳۹۰ با استفاده از ۵۸۴ باند فراطیفی بهترین ترکیبات باندی برای چهار شاخص SAVI2، NDWI، RWI و TSAVI به

مقادیر ریشه مربعات میانگین خطای بدست آمده با استفاده از روش Cross validation ( $RRMSE_{cv}$ ) در واقع همان خطای تست شبکه محسوب می شود چرا که هر ۹۵ نمونه تخمین زده شده هیچ یک در آموزش شبکه دخیل نبوده و در واقع هر یک نمونه تست می باشند. رابطه میان مقادیر پیش بینی شده و اندازه گیری شد با استفاده از ده مولفه اصلی به عنوان ورودی و بهترین ساختار شبکه در شکل ۳(b) نشان داده شده است.

این چهار شاخص با ترکیبات باندی بهینه که برای تخمین VWC بدست آمدند به عنوان ورودی برای شبکه عصبی مورد استفاده قرار گرفتند. الگوریتم های مورد استفاده برای آموزش شبکه، هنگامی که ورودی شاخص ها بودند در جدول ۳ نشان داده شده است. همانگونه که در این جدول نشان داده شده است، همانند شبکه با استفاده از مولفه به عنوان ورودی، در این حالت نیز شبکه با استفاده از الگوریتم آموزشی شیب توام (scg) در مقایسه با سایر الگوریتم های آموزشی مورد استفاده، دارای بهترین نتیجه بود (RRMSE<sub>cv</sub>=0.33, R<sup>2</sup><sub>cv</sub>=0.86). در الگوریتم آموزشی LM برای لایه پنهان ۶ و ۷ نرون مورد بررسی قرار گرفت و با افزایش یک نرون اضافی دقت تخمین کاهش یافت. رابطه میان مقادیر پیش بینی شده و اندازه گیری شده با استفاده از بهترین ساختار شبکه (هنگام استفاده از شاخص ها) در شکل (۳) نشان داده شده است.

منظور تخمین VWC انتخاب شدند [۱]. این ترکیبات باندی بهینه برای هر یک از شاخص ها بصورت زیر بود:

$$RWI = \frac{\rho_{720}}{\rho_{1418}} \quad (1)$$

$$NDWI = \frac{\rho_{679} - \rho_{1965}}{\rho_{679} + \rho_{1965}} \quad (2)$$

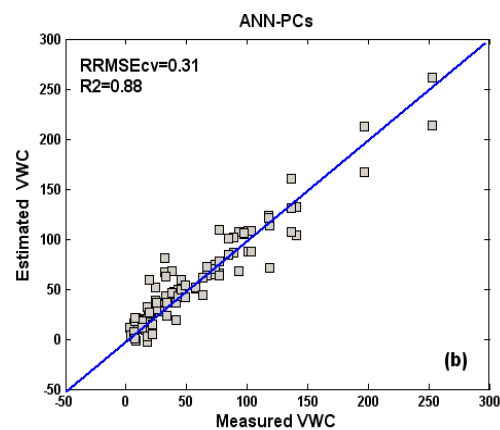
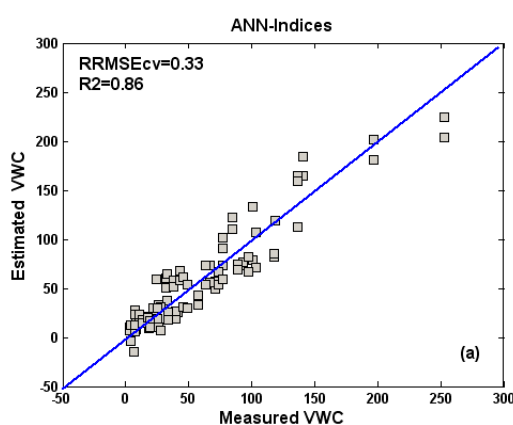
$$SAVI2 = \frac{\rho_{825}}{\rho_{1443} + \frac{a}{b}} \quad (3)$$

$$TSAVI = \frac{a(\rho_{2135} - a\rho_{2237} - b)}{a\rho_{2135} + \rho_{2237} - ab} \quad (4)$$

بطوری که  $\rho_i$  ها مقادیر بازتابندگی در طول موج  $i$  می باشد. برای مثال  $\rho_{720}$  مقدار بازتابندگی در طول موج ۷۲۰ نانومتر می باشد.  $a$  و  $b$  مقادیر بدست آمده برای خط خاک می باشند.

جدول ۳- شبکه های عصبی مختلف و نتایج بدست آمده از آنها در تخمین VWC با استفاده از شاخص ها باریک باند به عنوان ورودی

| ورودی ها                 | شبکه بهینه | تابع فعال سازی  | تابع آموزش | R <sup>2</sup> <sub>cv</sub> | RRMSE <sub>cv</sub> |
|--------------------------|------------|-----------------|------------|------------------------------|---------------------|
|                          | ۴*۶*۱      | {tansig,linear} | Oss        | ۰,۸۳                         | ۰,۳۶                |
| ۴ شاخص باریک باند بهینه: | ۴*۶*۱      | {tansig,linear} | Rp         | ۰,۸۵                         | ۰,۳۵                |
| RWI, NDWI, SAVI2 و       | ۴*۷*۱      | {tansig,linear} | Rp         | ۰,۷۸                         | ۰,۴۲                |
| TSAVI                    | ۴*۶*۱      | {tansig,linear} | Lm         | ۰,۸۴                         | ۰,۳۵                |
|                          | ۴*۶*۱      | {tansig,linear} | Scg        | ۰,۸۶                         | ۰,۳۳                |



شکل ۱- رابطه میان مقادیر پیش بینی شده و اندازه گیری شده VWC همراه با مقادیر RRMSE<sub>cv</sub> و R<sup>2</sup><sub>cv</sub> با استفاده از شبکه عصبی-

شاخص ها (a)، شبکه عصبی - مولفه های اصلی (b)

محتوای آبی گیاه و همچنین بهره بردن از تمامی باند های موجود در داده های فرایفی شد.

استفاده از شبکه با ساختار و پارامتر های مشابه اما ورودی های متفاوت نشان داد از جمله پارامتر های موثر

### ۳-۴ بحث

با استفاده از ورودی ها و ساختار های متفاوتی از شبکه پس انتشار خطا در این تحقیق سعی به برآورد

در نتیجه برآورد توسط شبکه های عصبی علاوه بر سایر پارامترها انتخاب ورودی مناسب می باشد. با استفاده از روش های کاهش ابعاد داده های ورودی که شامل استفاده از مولفه های اصلی به عنوان یک روش کاهش ابعاد داده رایج و همچنین استفاده از شاخص های باریک بند کارایی شبکه ها مورد مقایسه قرار گرفت. مولفه های اصلی دارای این مزیت می باشند که برخلاف باند های طیفی هیچ گونه همبستگی مابین آنها وجود ندارد [۲۱]. در مواردی که تعداد ورودی ها زیاد اما دارای همبستگی می باشند استفاده از مولفه های اصلی و همچنین شاخص ها بطور موثر توانستند ساختار شبکه را کوچک و محاسبات را کاهش دهند. استفاده از مولفه ها اغلب منجر به نتیجه بهتری در پیش بینی  $y$  می شوند. Ramadan و همکاران در سال ۲۰۰۵ برای برآورد پارامترهای خاک ۳۹ مولفه اصلی را به عنوان ورودی شبکه ارائه دادند که بهترین نتیجه را در مقایسه با سایر روش ها بدست آوردند. در تحقیق حاضر نیز استفاده از ۱۰ مولفه اصلی در مقایسه با شبکه با استفاده از تمامی باند ها و همچنین ۴ شاخص گیاهی به عنوان ورودی با تابع آموزشی یکسان، بهترین برآورد محتوای آبی گیاه را دارا بود:

$$(ANN-PC: R^2_{cv}=0.88, \quad RRMSE_{cv}=0.31, \\ ANN\_index: R^2_{cv}=0.86, \quad RRMSE_{cv}=0.33, \quad ANN: \\ R^2_{cv}=0.80, \quad RRMSE_{cv}=0.40)$$

همانگونه که مشاهده شد، تعداد نرون های لایه پنهان از جمله پارامترهای موثر در انتخاب یک شبکه مناسب (در این تحقیق تنها یک لایه پنهان) می باشد. تعداد بیش از حد کوچک نرون های لایه پنهان منجر به تخمین کمتر از مقدار واقعی<sup>۱</sup> و زیادی بیش از حد نرون ها منجر به بیش برآزش می شود [۱۵]. در تمامی شبکه های فوق مشاهده شد که افزایش تعداد نرون ها، موثر در بهبود تخمین بود اما افزایش بیش از حد منجر به بیش برآزش شبکه شد. این مشکل به دلایل متعددی از جمله افزایش بیش از ساختار شبکه، زیاد بودن تعداد پارامترهای ورودی نسبت به نمونه های آموزشی ایجاد می شود و بصورت دقت بالای داده های آموزشی و بسیار پایین نمونه های تست (در این تحقیق نمونه های Cross\_validate شده) قابل شناسایی

می باشد اگر چه تشخیص اندازه بهینه شبکه بسیار سخت می باشد. [۱۵]. از نمونه های بیش برآزش را می توان در تحقیق Ye و همکاران در سال ۲۰۰۶ برای برآورد میزان محصول که با استفاده از داده های فرا طیفی انجام شده بود مشاهده کرد. در تحقیق آنها که با استفاده از داده های سه ماهه انجام پذیرفته بود دقت داده های آموزشی بدست آمده برای دو ماه دارای اختلاف بسیاری با نمونه های تست ماه مربوطه داشت (دقت بالا برای نمونه های آموزشی و بسیار پایین برای نمونه های تست). این اختلاف نشان دهنده عدم تطبیق شبکه هنگام مواجهه با داده های جدید که به عبارت دیگر همان مشکل بیش برآزش بود، می باشد.

توابع آموزشی متفاوتی برای آموزش شبکه مورد آزمایش قرار گرفتند که تنها بهترین الگوریتم های مورد بررسی در جداول نشان داده شده اند. در بین توابع آموزشی ارائه شده الگوریتم شیب توام (TrainScg) نسبت به سایر الگوریتم ها در آموزش شبکه موفق تر بود. این الگوریتم تقریباً برای آموزش اکثر مسائل مناسب می باشد، خصوصاً شبکه های با مولفه های وزن و بایاس بسیار زیاد این الگوریتم مناسبترین می باشد [۱۵]. الگوریتم پس انتشار ارتجاعی (Trainrp) از دیگر الگوریتم هایی است که برای شبکه های بزرگ مناسب می باشد. و همانطور که در جداول مشاهده می شود برای شبکه ها با ورودی مولفه ها و شاخص ها پس از الگوریتم TrainScg دارای بهترین تخمین بود.

$$(ANN-PC: R^2_{cv}=0.86, \quad RRMSE_{cv}=0.33, \\ ANN\_index: R^2_{cv}=0.85, \quad RRMSE_{cv}=0.35)$$

#### ۴- نتیجه گیری

توسعه در زمینه سنجش از دور فراطیفی و تصویر برداری اسپکترومتری دریچه ای جدید به سمت پایش و تخمین پارامترهای گیاهی گشوده است. استفاده از شبکه های عصبی در سنجش از دور اغلب در طبقه بندی مورد استفاده قرار می گیرد. در این تحقیق که بر اساس داده های فراطیفی بدست آمده با استفاده از اندازه گیری های آزمایشگاهی انجام شد توانایی شبکه های برای تخمین محتوای آبی گیاهان با استفاده از داده های

<sup>۱</sup> Underestimate



الگوریتم آموزشی مناسب برای آن انتخاب شد، مدلی با توانایی بالا در تخمین محتوای آبی با استفاده از داده های فراطیفی می باشند و نهایتاً در این تحقیق داد شبکه عصبی با استفاده از مولفه های اصلی والگوریتم آموزشی شیب توام (scg) دارای بهترین برآورد میزان VWC بود.

فراطیفی مورد بررسی قرار گرفت. شبکه عصبی پس انتشار با الگوریتم های آموزشی و ساختار متفاوت برای تخمین VWC مورد استفاده قرار گرفت که نتایج نشان داد انتخاب معماری شبکه در آموزش بهتر شبکه و همچنین جلوگیری از مشکل بیش برآزش بسیار موثر می باشد. شبکه های عصبی مصنوعی، هنگامی که شبکه بهینه با ساختار و

## مراجع

- [۱] مژگان میرزایی، روشنک درویش زاده، علیرضا شکیبیا، علی اکبر متکان (۱۳۹۰) انتخاب شاخص های فراطیفی (باریک باند) بهینه برای تخمین محتوای آبی گیاهان با در نظر گرفتن شرایط متفاوت تراکم تاج پوشش گیاه و خاک پس زمینه، نشریه علمی- پژوهشی سنجش از دور و GIS ایران. سال سوم. شماره اول، بهار ۱۳۹۰. ۷۰-۵۵.
- [2] Banskota, A. (2006). Estimating leaf area index of salt marsh vegetation using airborne hyperspectral data.
- [3] Broge, N., & Leblanc, E. (2001). Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment* 76 (2), 156–172. , 156–172.
- [4] Carter, G.A., 1994. Ratios of leaf reflectances in narrow wavebands as indicators of plant stress. *International Journal of Remote Sensing*, 15(3): 697-703.
- [5] Ceccato, P., Flasse, S., Tarantola, S., Jacquemoud, S., & Gregoire, J.-M. (2000). Detecting vegetation leaf water content using reflectance in the optical domain. *Remote Sensing of Environment* , 77 , 22– 33.
- [6] Ceccato, P., Gobron, N., Flasse, S., Pinty, B., & Tarantola, S. (2002). Designing a spectral index to estimate vegetation water content from remote sensing data: Part 1 Theoretical approach. *Remote Sensing of Environment* , 82, 188–197.
- [7] Chang, D., & Islam, S. (2000). Estimation of soil physical properties using remote sensing and artificial neural network. *Remote Sensing of Environment* , 74 (3), 534–544.
- [8] Chen, D., Huang, J., & Jackson, T. J. (2005). Vegetation water content estimation for corn and soybeans using spectral indices derived from MODIS near- and short-wave infrared bands. *Remote Sensing of Environment* , 98, 225 – 236
- [9] Cho, M. A. (2007). *Hyperspectral remote sensing of biochemical and biophysical parameters*. International Institute for Geo-information Science & Earth Observation, Enschede, the Netherlands (ITC).
- [10] Clevers, J., kooistra, L., & Schaepman, M. (2008). Using spectral information from the NIR water absorption features for retrieval of canopy water content. *International journal of Applied Earth Observation and Geoinformation* , 10, 388-397.
- [11] Clevers, J., Kooistra, L., & Schaepma, M. (2010). Estimating canopy water content using hyperspectral remote sensing data. *International Journal of Applied Earth Observation and Geoinformation* , 12, 119–125.
- [12] Darvishzadeh, R. (2008). *Hyperspectral remote sensing of vegetation parameters using statistical and physical models*. International Institute for Geo-information Science & Earth Observation, Enschede, the Netherlands (ITC).
- [13] Darvishzadeh, R., Skidmore, A., Schlerf, M., Atzberger, C., Corsi, F., & Cho, M. (2008 ). LAI and chlorophyll estimation for a heterogeneous grassland using hyperspectral measurements. *ISPRS Journal of Photogrammetry & Remote Sensing* , 63, 409–426.

- [14] Del Frate, F., Ferrazzoli, P., & Schiavon, G. (2003). Retrieving soil moisture and agricultural variables by microwave radiometry using neural networks. *Remote Sensing of Environment* , 84 (2), 174–183.
- [15] Demuth, H., Beale, M., & Hagan, M. (2010). *Neural Network Toolbox™ 6 User's guide*.
- [16] Drummond, S., Sudduth, K., Joshi, A., Birrell, S., & Kitchen, N. (2003). Statistical and neural methods for site-specific yield prediction. *Trans. ASAE* , 46 (1), 5–14.
- [17] Jackson, T. J., Chen, D., Cosh, M., Li, F., Anderson, M., Walthall, C., et al. (2003). Vegetation water content mapping using Landsat data derived normalized difference water index for corn and soybeans. *Remote Sensing of Environment* , 92 (475–482).
- [18] Jin, Y., & Liu, C. (1997). Biomass retrieval from high-dimensional active/passive remote sensing data by using artificial neural networks. *Int. J. Remote Sens* , 18 (4), 971–979.
- [19] Liu, M., Liu, X., Li, M., Fang, M., & Chi, W. (2010). Neural-network model for estimating leaf chlorophyll concentration in rice under stress from heavy metals using four spectral indices. *biosystems engineering* , 223-233.
- [20] Pyne, S. J., Andrews, P. L., & Laven, R. D. (1996). *Introduction to wildland fire*. New York: Wiley.
- [21] Ramadan, Z., Hopke, P. K., Johnson, M. J., & Scow, K. M. (2005). Application of PLS and Back-Propagation Neural Networks for the estimation of soil properties. *Chemometrics and Intelligent Laboratory Systems* , 75 , 23– 30.
- [22] Sapuan, S., & Mujtaba, I. (2010). *Composit Materials Technology, Neural Network Applications*. Taylor and Francis.
- [23] Savitzky, A., & Golay, M. (1964). Smoothing and differentiation of data by simplified least square procedure.
- [24] Suárez, L., Zarco-Tejada, P., & Berni, J. (2009). Modelling PRI for water stress detection using radiative transfer models. *Remote Sensing of Environment* , 113,730–744.
- [25] Tso, B., & Mather, P. (2001). *classification methods for remotely sensed data*. Talor & Francis.
- [26] Uno, Y., Prasher, S., Lacroix, R., Goel, P., Karimi, Y., Viau, A., et al. (2005). Artificial neural networks to predict corn yield from Compact Airborne Spectrographic Imager data. *Computers and Electronics in Agriculture* , 47,149-161.
- [27] Ye, X., Sakai, K., Garciano, L. O., Asada, S.-I., & Sasao, A. (2006). Estimation of citrus yield from airborne hyperspectral images using a neural network model. *ecological modelling* , 198,426-432.
- [28] zarco-Tejada, P., Millerb, J., Moralese, A., Berjo'na, A., & Agu'era, J. (2004). Hyperspectral indices and model simulation for chlorophyll estimation in open-canopy tree crops. *Remote Sensing of Environment* , 90, 463–476.