

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده مهندسی عمران

پایان نامه کارشناسی ارشد مهندسی آب و سازه های هیدرولیکی

پیش بینی غلظت سختی موقت و دائم در آب زیرزمینی براساس پارامترهای
شیمیایی با ترکیب شبکه عصبی و سیستم اطلاعات مکانی جغرافیایی

نگارنده: سید مصطفی موسوی

اساتید راهنما

دکتر احمد احمدی

دکتر فرهاد قادری

شهریور ۱۴۰۰

در این صفحه صورت جلسه دفاع را قرار دهید. لازم است پس از صحافی این صفحه مجدداً توسط دانشکده مهر گردد و استاد راهنما با امضای خود اصلاحات پایان نامه را تایید کند.

تقديم به

رنگان بی برکت...

تهدنامه

اینجانب سید مصطفی موسوی دانشجوی دوره کارشناسی ارشد رشته آب و سازه‌های هیدرولیکی دانشکده عمران دانشگاه صنعتی شاهرود نویسنده پایان‌نامه پیش بینی غلظت سختی موقت و دائم در آب زیرزمینی بر اساس پارامترهای شیمیایی با ترکیب شبکه عصبی و سیستم اطلاعات مکانی جغرافیایی (GIS) تحت راهنمایی دکتر احمد احمدی-دکتر فرهاد قادری متعهد می‌شوم.

- تحقیقات در این پایان‌نامه توسط اینجانب انجام شده است و از صحت و اصال برخوردار است .
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است .
- مطالب مندرج در پایان‌نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است .
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می‌باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید .
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان‌نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان‌نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان‌نامه ، در مواردی که از موجود زنده (یا بافت‌های آن‌ها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است .
- در کلیه مراحل انجام این پایان‌نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری ، ضوابط و اصول اخلاق انسانی رعایت شده است .

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج ، کتاب ، برنامه های رایانه ای ، نرم افزار ها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می‌باشد . این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود . استفاده از اطلاعات و نتایج موجود در پایان‌نامه بدون ذکر مرجع مجاز نمی‌باشد.

چکیده

در صورتی که منبع دیگری برای تامین آب در دسترس نباشد، آب‌های زیرزمینی تنها منبع در بسیاری از مناطق است. کاهش آب‌های زیرزمینی، آلودگی و عدم مدیریت مناسب آن مسائل انسانی را تحت تاثیر قرار داده است. با توجه به تاثیر آب بر سلامت انسان، اهمیت در نظر گرفتن کیفیت آن نیازمند تحقیقاتی است که توانایی برنامه‌ریزی برای کنترل کیفیت را دارد. در این زمینه مدل‌های هوشمند و داده محور از جمله روش‌هایی است که با توجه به سهولت انجام آن نسبت به سایر روش‌ها به سرعت در حال گسترش هستند. این مدل‌ها توانایی یادگیری و تعمیم داده‌ها را داشته و قادر به حل مسائل مربوط به پیش‌بینی، مدیریت و تخمین در جنبه‌های مختلفی از منابع آب را دارند. در این میان شبکه عصبی ابرازی قدرتمند در حل مسائل رگرسیون و دسته‌بندی می‌باشد که توانایی حل مسائل با هیدرولوژی پیچیده با داده‌های کم، نسبت به سایر روش‌های داده محور را دارد. در این مطالعه، روش شبکه عصبی پیشرو و ماشین بردار پشتیبان جهت پیش‌بینی غلظت سختی موقت و دائم دشت آمل-بابل با استفاده از پارامترهای کیفی مانند سولفات، کلرید، نیترات، پی‌اچ، نسبت جذب سدیم، سختی کل، هدایت الکتریکی و پتاسیم به کار گرفته شد. برای این منظور از داده‌های کیفی جمع آوری شده از ۸۵۵ چاه در دو فصل پاییز و بهار به مدت شش سال استفاده شده است. سپس با ایجاد مدل شبکه عصبی پیشرو و مدل ماشین بردار پشتیبان با پارامترهای متفاوت، مدل بهینه انتخاب و مورد آموزش و آزمون قرار گرفت. نتایج بدست آمده براساس معیارهای R^2 ، MSE و RMSE برای شبکه عصبی پیشرو در سختی دائم به ترتیب ۰/۹۳۰۳، ۷۴۸/۷۶ و ۲۷/۳۶ و برای سختی موقت برابر ۰/۹۶۸۸، ۵۰۴/۷۸ و ۲۲/۴۶ می‌باشد. این نتایج برای ماشین بردار پشتیبان در سختی دائم برابر ۰/۹۳۰۴، ۰/۰۰۰۱۴۰۱، ۰/۰۱۸۱ و برای سختی موقت برابر ۰/۹۸۳۳، ۰/۰۰۰۰۳۳۸، ۰/۰۰۵۸ می‌باشد.

نتایج بدست آمده از مدلسازی سختی موقت و دائم در دشت آمل-بابل نشان دهنده توانایی مدل‌های ایجاد شده در مدلسازی بوده و با تقریب مناسبی میزان غلظت سختی را در منطقه مورد نظر پیش‌بینی کرده است. غلظت سختی در قسمت شرق و مرکزی نسبت به سایر نقاط بیشتر بوده بطوری که از سمت شرق به مرکز حرکت می‌کند.

کلمات کلیدی: شبکه عصبی، ماشین بردار پشتیبان، سختی موقت، سختی دائم، آب‌های زیرزمینی

لیست مقالات مستخرج از پایان نامه

۱-

۲-

۳-

فهرست مطالب

ر	فهرست جداول
ز	فهرست اشکال
۱	فصل ۱: مقدمه
۲	۱-۱ اهمیت موضوع.....
۴	۱-۲ اهداف تحقیق.....
۵	۱-۳ روش انجام تحقیق.....
۵	۱-۴ ساختار پایان نامه.....
۷	فصل ۲: مبانی و روش تحقیق
۷	۲-۱ مقدمه.....
۸	۲-۲ مدل های هیدرولوژی.....
۱۱	۲-۳ شبکه های عصبی مصنوعی (ANN).....
۱۲	۲-۳-۱ ساختار شبکه عصبی.....
۱۳	۲-۳-۲ آموزش شبکه.....
۱۴	۲-۳-۳ روش های یادگیری.....
۱۵	۲-۳-۴ یادگیری نظارت شده.....
۱۶	۲-۳-۵ یادگیری بدون نظارت.....
۱۶	۲-۳-۶ پرسپترون چندلایه (MLP).....
۱۹	۲-۳-۷ شبکه عصبی پیشرو.....
۲۱	۲-۳-۸ الگوریتم پسرو.....
۲۲	۲-۴ تابع فعالساز.....

۲۳	 Linear تابع ۲-۴-۱
۲۳	 Sigmoid تابع ۲-۴-۲
۲۴	 Tanh تابع ۲-۴-۳
۲۵	 ماشین بردار پشتیبان ۲-۵
۲۶	 دسته بندی بردار پشتیبان ۲-۶
۲۶	 رگرسیون بردار پشتیبان ۲-۷
۲۷	 حالت خطی رگرسیون بردار پشتیبان ۲-۷-۱
۳۰	 حالت غیرخطی رگرسیون بردار پشتیبان ۲-۷-۲
۳۱	 سختی آب ۲-۸

فصل ۳: مبانی نظری و پیشینه تحقیق ۳۳

۳۳	 مقدمه ۳-۱
۳۵	 مروری بر چند مطالعه در زمینه مدلسازی کیفی به کمک شبکه عصبی ۳-۲
۴۱	 منطقه مورد مطالعه ۳-۳

فصل ۴: نتایج و بحث ۴۳

۴۳	 مقدمه ۴-۱
۴۴	 مشخصات پارامترهای ورودی و خروجی ۴-۲
۴۷	 سولفات (Sulfate) ۴-۲-۱
۴۸	 کلرید (Chloride) ۴-۲-۲
۴۹	 نیترات (Nitrate) ۴-۲-۳
۵۰	 پی اچ (PH) ۴-۲-۴
۵۱	 سختی کل (TH) ۴-۲-۵
۵۲	 نسبت جذب سدیم (SAR) ۴-۲-۶

۵۳	۴-۲-۷ هدایت الکتریکی (EC)
۵۴	۴-۲-۸ پتاسیم (Potassium)
۵۵	۴-۲-۹ سختی موقت و دائم آب
۵۶	۴-۳ پهنه بندی در GIS و استخراج داده ها
۸۰	۴-۴ معیار عملکرد
۸۱	۴-۵ مدل سازی به روش شبکه عصبی پیشرو
۸۱	۴-۶ تعیین الگوریتم آموزش و یادگیری
۸۲	۴-۷ آموزش و تحلیل شبکه
۸۵	۴-۸ پیش بینی سختی موقت و دائم آب به کمک شبکه عصبی پیشرو
۹۱	۴-۹ مدل سازی به روش رگرسیون بردار پشتیبان
۹۲	۴-۱۰ نرمال سازی داده ها
۹۲	۴-۱۱ تعیین تابع کرنل
۹۳	۴-۱۲ مدل سازی سختی موقت و دائم به کمک SVM
۹۶	۴-۱۳ پیش بینی سختی موقت و دائم به کمک SVM
۱۰۲	۴-۱۴ تحلیل حساسیت

فصل ۵: نتایج و پیشنهادات

۱۰۴	۵-۱ نتایج
۱۰۶	۵-۲ پیشنهادات

مراجع

۱۰۷

فهرست جداول

جدول ۱-۲. برخی از توابع کرنل پر کاربرد.....	۳۰
جدول ۱-۴. اطلاعات آماری کیفیت آب های زیرزمینی.....	۴۴
جدول ۲-۴. ضریب همبستگی خطی (R) بین سختی دائم و سختی موقت و پارامترهای ورودی ...	۴۶
جدول ۳-۴. نتایج شبکه عصبی Feedforward سختی دائم.....	۸۵
جدول ۴-۴. نتایج شبکه عصبی Feedforward سختی موقت.....	۸۵
جدول ۵-۴. نتایج مدل Support Vector سختی دائم.....	۹۳
جدول ۶-۴. نتایج مدل Support Vector سختی موقت.....	۹۳

فهرست اشکال

شکل ۱-۲. اتصالات نورو	۱۲
شکل ۲-۲. معماری پرسپترون تک لایه	۱۷
شکل ۳-۲. ساختار پرسپترون چند لایه	۱۸
شکل ۴-۲. نمونه‌ای از شبکه عصبی پیشرو	۲۰
شکل ۵-۲. نمونه‌ای از شبکه عصبی پسرو	۲۲
شکل ۶-۲. تابع سیگموئید	۲۳
شکل ۷-۲. تابع تانژانت هیپربولیک	۲۴
شکل ۸-۲. ابرصفحه جداکننده بهینه	۲۶
شکل ۹-۲. توابع خطا	۲۷
شکل ۱۰-۲. تابع خطای ماشین بردار پشتیبان خطی	۲۸
شکل ۱-۳. نقشه کاربری اراضی بابل-آمل	۴۱
شکل ۲-۳. نقشه ضخامت آبخوان	۴۲
شکل ۱-۴. هیستوگرام فراوانی داده‌های سولفات	۴۷
شکل ۲-۴. هیستوگرام فراوانی داده‌های کلرید	۴۸
شکل ۳-۴. هیستوگرام فراوانی داده‌های نترات	۴۹
شکل ۴-۴. هیستوگرام فراوانی داده‌های PH	۵۰
شکل ۶-۴. هیستوگرام فراوانی داده‌های سختی کل	۵۱
شکل ۷-۴. هیستوگرام فراوانی داده‌های SAR	۵۲
شکل ۸-۴. هیستوگرام فراوانی داده‌های EC	۵۳
شکل ۹-۴. هیستوگرام فراوانی داده‌های K	۵۴
شکل ۱۳-۴. هیستوگرام فراوانی داده‌های سختی دائم	۵۵
شکل ۱۴-۴. هیستوگرام فراوانی داده‌های سختی موقت	۵۶
شکل ۱۱-۴. سولفات فصل پاییز ۹۲	۵۷
شکل ۱۲-۴. سولفات فصل بهار ۹۳	۵۷
شکل ۱۳-۴. سولفات فصل پاییز ۹۳	۵۷
شکل ۱۴-۴. سولفات فصل بهار ۹۴	۵۷
شکل ۱۵-۴. سولفات فصل پاییز ۹۴	۵۷
شکل ۱۶-۴. سولفات فصل بهار ۹۵	۵۷
شکل ۱۷-۴. سولفات فصل پاییز ۹۵	۵۸
شکل ۱۸-۴. سولفات فصل بهار ۹۶	۵۸

شکل ۴-۱۹. سولفات فصل پاییز ۹۶.....	۵۸
شکل ۴-۲۰. سولفات فصل بهار ۹۷.....	۵۸
شکل ۴-۲۱. سولفات فصل پاییز ۹۷.....	۵۸
شکل ۴-۲۲. سولفات فصل بهار ۹۸.....	۵۸
شکل ۴-۲۳. سولفات فصل پاییز ۹۸.....	۵۹
شکل ۴-۲۴. کلر فصل پاییز ۹۲.....	۶۰
شکل ۴-۲۵. کلر فصل بهار ۹۳.....	۶۰
شکل ۴-۲۶. کلر فصل پاییز ۹۳.....	۶۰
شکل ۴-۲۷. کلر فصل بهار ۹۴.....	۶۰
شکل ۴-۲۸. کلر فصل پاییز ۹۴.....	۶۰
شکل ۴-۲۹. کلر فصل بهار ۹۵.....	۶۰
شکل ۴-۳۰. کلر فصل پاییز ۹۵.....	۶۱
شکل ۴-۳۱. کلر فصل بهار ۹۶.....	۶۱
شکل ۴-۳۲. کلر فصل پاییز ۹۶.....	۶۱
شکل ۴-۳۳. کلر فصل بهار ۹۷.....	۶۱
شکل ۴-۳۴. کلر فصل پاییز ۹۷.....	۶۱
شکل ۴-۳۵. کلر فصل بهار ۹۸.....	۶۱
شکل ۴-۳۶. کلر فصل پاییز ۹۸.....	۶۲
شکل ۴-۳۷. نیترا ت فصل پاییز ۹۲.....	۶۳
شکل ۴-۳۸. نیترا ت فصل بهار ۹۳.....	۶۳
شکل ۴-۳۹. نیترا ت فصل پاییز ۹۳.....	۶۳
شکل ۴-۴۰. نیترا ت فصل بهار ۹۴.....	۶۳
شکل ۴-۴۱. نیترا ت فصل پاییز ۹۴.....	۶۳
شکل ۴-۴۲. نیترا ت فصل بهار ۹۵.....	۶۳
شکل ۴-۴۳. نیترا ت فصل پاییز ۹۵.....	۶۴
شکل ۴-۴۴. نیترا ت فصل بهار ۹۶.....	۶۴
شکل ۴-۴۵. نیترا ت فصل پاییز ۹۶.....	۶۴
شکل ۴-۴۶. نیترا ت فصل بهار ۹۷.....	۶۴
شکل ۴-۴۷. نیترا ت فصل پاییز ۹۷.....	۶۴
شکل ۴-۴۸. نیترا ت فصل بهار ۹۸.....	۶۴
شکل ۴-۴۹. نیترا ت فصل پاییز ۹۸.....	۶۵
شکل ۴-۵۰. پی اچ فصل پاییز ۹۲.....	۶۶

شکل ۴-۵۱. پی اچ فصل بهار ۹۳.....	۶۶
شکل ۴-۵۲. پی اچ فصل پاییز ۹۳.....	۶۶
شکل ۴-۵۳. پی اچ فصل بهار ۹۴.....	۶۶
شکل ۴-۵۴. پی اچ فصل پاییز ۹۴.....	۶۶
شکل ۴-۵۵. پی اچ فصل بهار ۹۵.....	۶۶
شکل ۴-۵۶. پی اچ فصل پاییز ۹۵.....	۶۷
شکل ۴-۵۷. پی اچ فصل بهار ۹۶.....	۶۷
شکل ۴-۵۸. پی اچ فصل پاییز ۹۶.....	۶۷
شکل ۴-۵۹. پی اچ فصل بهار ۹۷.....	۶۷
شکل ۴-۶۰. پی اچ فصل پاییز ۹۷.....	۶۷
شکل ۴-۶۱. پی اچ فصل بهار ۹۸.....	۶۷
شکل ۴-۶۲. پی اچ فصل پاییز ۹۸.....	۶۸
شکل ۴-۶۳. سختی کل فصل پاییز ۹۲.....	۶۹
شکل ۴-۶۴. سختی کل فصل بهار ۹۳.....	۶۹
شکل ۴-۶۵. سختی کل فصل پاییز ۹۳.....	۶۹
شکل ۴-۶۶. سختی کل فصل بهار ۹۴.....	۶۹
شکل ۴-۶۷. سختی کل فصل پاییز ۹۴.....	۶۹
شکل ۴-۶۸. سختی کل فصل بهار ۹۵.....	۶۹
شکل ۴-۶۹. سختی کل فصل پاییز ۹۵.....	۷۰
شکل ۴-۷۰. سختی کل فصل بهار ۹۶.....	۷۰
شکل ۴-۷۱. سختی کل فصل پاییز ۹۶.....	۷۰
شکل ۴-۷۲. سختی کل فصل بهار ۹۷.....	۷۰
شکل ۴-۷۳. سختی کل فصل پاییز ۹۷.....	۷۰
شکل ۴-۷۴. سختی کل فصل بهار ۹۸.....	۷۰
شکل ۴-۷۵. سختی کل فصل پاییز ۹۸.....	۷۱
شکل ۴-۷۶. نسبت جذب سدیم فصل پاییز ۹۲.....	۷۲
شکل ۴-۷۷. نسبت جذب سدیم فصل بهار ۹۳.....	۷۲
شکل ۴-۷۸. نسبت جذب سدیم فصل پاییز ۹۳.....	۷۲
شکل ۴-۷۹. نسبت جذب سدیم فصل بهار ۹۴.....	۷۲
شکل ۴-۸۰. نسبت جذب سدیم فصل پاییز ۹۴.....	۷۲
شکل ۴-۸۱. نسبت جذب سدیم فصل بهار ۹۵.....	۷۲
شکل ۴-۸۲. نسبت جذب سدیم فصل پاییز ۹۵.....	۷۳

شکل ۴-۸۳. نسبت جذب سدیم فصل بهار ۹۶.....	۷۳
شکل ۴-۸۴. نسبت جذب سدیم فصل پاییز ۹۶.....	۷۳
شکل ۴-۸۵. نسبت جذب سدیم فصل بهار ۹۷.....	۷۳
شکل ۴-۸۶. نسبت جذب سدیم فصل پاییز ۹۷.....	۷۳
شکل ۴-۸۷. نسبت جذب سدیم فصل بهار ۹۸.....	۷۳
شکل ۴-۸۸. نسبت جذب سدیم فصل پاییز ۹۸.....	۷۴
شکل ۴-۸۹. هدایت الکتریکی فصل پاییز ۹۲.....	۷۵
شکل ۴-۹۰. هدایت الکتریکی فصل بهار ۹۳.....	۷۵
شکل ۴-۹۱. هدایت الکتریکی فصل پاییز ۹۳.....	۷۵
شکل ۴-۹۲. هدایت الکتریکی فصل بهار ۹۴.....	۷۵
شکل ۴-۹۳. هدایت الکتریکی فصل پاییز ۹۴.....	۷۵
شکل ۴-۹۴. هدایت الکتریکی فصل بهار ۹۵.....	۷۵
شکل ۴-۹۵. هدایت الکتریکی فصل پاییز ۹۵.....	۷۶
شکل ۴-۹۶. هدایت الکتریکی فصل بهار ۹۶.....	۷۶
شکل ۴-۹۷. هدایت الکتریکی فصل پاییز ۹۶.....	۷۶
شکل ۴-۹۸. هدایت الکتریکی فصل بهار ۹۷.....	۷۶
شکل ۴-۹۹. هدایت الکتریکی فصل پاییز ۹۷.....	۷۶
شکل ۴-۱۰۰. هدایت الکتریکی فصل بهار ۹۸.....	۷۶
شکل ۴-۱۰۱. هدایت الکتریکی فصل پاییز ۹۸.....	۷۷
شکل ۴-۱۰۲. پتاسیم فصل پاییز ۹۲.....	۷۸
شکل ۴-۱۰۳. پتاسیم فصل بهار ۹۳.....	۷۸
شکل ۴-۱۰۴. پتاسیم فصل پاییز ۹۳.....	۷۸
شکل ۴-۱۰۵. پتاسیم فصل بهار ۹۴.....	۷۸
شکل ۴-۱۰۶. پتاسیم فصل پاییز ۹۴.....	۷۸
شکل ۴-۱۰۷. پتاسیم فصل بهار ۹۵.....	۷۸
شکل ۴-۱۰۸. پتاسیم فصل پاییز ۹۵.....	۷۹
شکل ۴-۱۰۹. پتاسیم فصل بهار ۹۶.....	۷۹
شکل ۴-۱۱۰. پتاسیم فصل پاییز ۹۶.....	۷۹
شکل ۴-۱۱۱. پتاسیم فصل بهار ۹۷.....	۷۹
شکل ۴-۱۱۲. پتاسیم فصل پاییز ۹۷.....	۷۹
شکل ۴-۱۱۳. پتاسیم فصل بهار ۹۸.....	۷۹
شکل ۴-۱۱۴. پتاسیم فصل پاییز ۹۸.....	۸۰

- شکل ۴-۱۱۵. (الف) نمودار عملکرد فرایند آموزش سختی دائم برای رسیدن به کمترین خطا ۸۲
- شکل ۴-۱۱۶. (الف) نمودار رگرسیون بین داده‌های محاسباتی و مشاهداتی به ازای کل داده‌های سختی دائم ۸۳
- شکل ۴-۱۱۷. (الف) مقایسه غلظت سختی دائم پیش بینی شده و اندازه‌گیری در مرحله آزمون مدل ANN ۸۴
- شکل ۴-۱۱۸. سختی دائم پاییز ۹۲ ۸۶
- شکل ۴-۱۱۹. سختی دائم بهار ۹۳ ۸۶
- شکل ۴-۱۲۰. سختی دائم پاییز ۹۳ ۸۶
- شکل ۴-۱۲۱. سختی دائم بهار ۹۴ ۸۶
- شکل ۴-۱۲۲. سختی دائم پاییز ۹۴ ۸۶
- شکل ۴-۱۲۳. سختی دائم بهار ۹۵ ۸۶
- شکل ۴-۱۲۴. سختی دائم پاییز ۹۵ ۸۷
- شکل ۴-۱۲۵. سختی دائم بهار ۹۶ ۸۷
- شکل ۴-۱۲۶. سختی دائم پاییز ۹۶ ۸۷
- شکل ۴-۱۲۷. سختی دائم بهار ۹۷ ۸۷
- شکل ۴-۱۲۸. سختی دائم پاییز ۹۷ ۸۷
- شکل ۴-۱۲۹. سختی دائم بهار ۹۸ ۸۷
- شکل ۴-۱۳۰. سختی دائم پاییز ۹۸ ۸۸
- شکل ۴-۱۳۱. سختی موقت پاییز ۹۲ ۸۹
- شکل ۴-۱۳۲. سختی موقت بهار ۹۳ ۸۹
- شکل ۴-۱۳۳. سختی موقت پاییز ۹۳ ۸۹
- شکل ۴-۱۳۴. سختی موقت بهار ۹۴ ۸۹
- شکل ۴-۱۳۵. سختی موقت پاییز ۹۴ ۸۹
- شکل ۴-۱۳۶. سختی موقت بهار ۹۵ ۸۹
- شکل ۴-۱۳۷. سختی موقت پاییز ۹۵ ۹۰
- شکل ۴-۱۳۸. سختی موقت بهار ۹۶ ۹۰
- شکل ۴-۱۳۹. سختی موقت پاییز ۹۶ ۹۰
- شکل ۴-۱۴۰. سختی موقت بهار ۹۷ ۹۰
- شکل ۴-۱۴۱. سختی موقت پاییز ۹۷ ۹۰
- شکل ۴-۱۴۲. سختی موقت بهار ۹۸ ۹۰
- شکل ۴-۱۴۳. سختی موقت پاییز ۹۸ ۹۱
- شکل ۴-۱۴۴. (الف) مقایسه سختی دائم پیش بینی شده و اندازه‌گیری در مرحله آموزش ۹۴

شکل ۴-۱۴۵. (الف) مقایسه سختی موقت پیش بینی شده و اندازه گیری در مرحله آموزش	۹۵
شکل ۴-۱۴۶. سختی دائم پاییز ۹۲.....	۹۶
شکل ۴-۱۴۷. سختی دائم بهار ۹۳	۹۶
شکل ۴-۱۴۸. سختی دائم پاییز ۹۳	۹۷
شکل ۴-۱۴۹. سختی دائم بهار ۹۴	۹۷
شکل ۴-۱۵۰. سختی دائم پاییز ۹۴	۹۷
شکل ۴-۱۵۱. سختی دائم بهار ۹۵	۹۷
شکل ۴-۱۵۲. سختی دائم پاییز ۹۵	۹۷
شکل ۴-۱۵۳. سختی دائم بهار ۹۶	۹۷
شکل ۴-۱۵۴. سختی دائم پاییز ۹۶	۹۸
شکل ۴-۱۵۵. سختی دائم بهار ۹۷	۹۸
شکل ۴-۱۵۶. سختی دائم پاییز ۹۷	۹۸
شکل ۴-۱۵۷. سختی دائم بهار ۹۸	۹۸
شکل ۴-۱۵۸. سختی دائم پاییز ۹۸	۹۸
شکل ۴-۱۵۹. سختی موقت پاییز ۹۲	۹۹
شکل ۴-۱۶۰. سختی موقت بهار ۹۳	۹۹
شکل ۴-۱۶۱. سختی موقت پاییز ۹۳	۹۹
شکل ۴-۱۶۲. سختی موقت بهار ۹۴	۹۹
شکل ۴-۱۶۳. سختی موقت پاییز ۹۴	۱۰۰
شکل ۴-۱۶۴. سختی موقت بهار ۹۵	۱۰۰
شکل ۴-۱۶۵. سختی موقت پاییز ۹۵	۱۰۰
شکل ۴-۱۶۶. سختی موقت بهار ۹۶	۱۰۰
شکل ۴-۱۶۷. سختی موقت پاییز ۹۶	۱۰۰
شکل ۴-۱۶۸. سختی موقت بهار ۹۷	۱۰۰
شکل ۴-۱۶۹. سختی موقت پاییز ۹۷	۱۰۱
شکل ۴-۱۷۰. سختی موقت بهار ۹۸	۱۰۱
شکل ۴-۱۷۱. سختی موقت پاییز ۹۸	۱۰۱
شکل ۴-۱۷۲. نتایج بدست آمده از تحلیل حساسیت سختی دائم و موقت.....	۱۰۳

فصل ۱ : مقدمه

۱-۱ اهمیت موضوع

در ایران میانگین متوسط سالانه میزان بارندگی به دلیل شرایط آب و هوایی ۲۵۰ میلی‌متر است که در مقایسه با میانگین متوسط بارندگی در سطح خشکی که ۸۶۰ میلی‌متر تخمین زده شده است، ملاحظه خواهد شد که بارندگی در ایران حتی کمتر از یک سوم متوسط بارندگی در سطح دنیاست. علاوه بر این زمان ریزش نزولات جوی و محل ریزش آن‌ها نیز با نیاز بخش کشاورزی، که مصرف کننده اصلی آب در کشور می باشد، مطابقت ندارد. اکثر شهرهای ایران در مناطقی واقع اند که به رودخانه‌هایی که جریان آب آن‌ها مستقیماً از رواناب حاصل از بارندگی‌ها تامین شده باشد دسترسی ندارند. بنابراین باید پذیرفت که خشکی در ایران یک واقعیت اقلیمی است و این ما هستیم که باید خود را با آن سازگاری دهیم. معیاری که برای پرآبی یا کم آبی یک کشور بکار برده می‌شود سرانه آب قابل تجدید در آن کشور است. در حال حاضر با توجه به جمعیت جهان مقدار آب تجدید شونده دنیا حدود ۶۵۰۰ مترمکعب در سال برای هر نفر است [۱].

سختی، توسط انواع یون‌های فلزی چند ظرفیتی محلول که عمدتاً توسط کاتیون‌های کلسیم و منیزیم ایجاد می‌شود. اگرچه کاتیون‌های دیگر (به عنوان مثال آلومینیوم، باریم، آهن، منگنز، استرانسیم و روی) نیز در آن نقش دارند. به طور کلی آب حاوی کربنات کلسیم در غلظت های زیر ۶۰ میلی‌گرم در لیتر نرم، ۶۰ تا ۱۲۰ میلی‌گرم در لیتر سختی متوسط، ۱۲۰ تا ۱۸۰ میلی‌گرم در لیتر سخت و بیشتر از ۱۸۰ میلی‌گرم در لیتر بسیار سخت تلقی می‌شود. اگرچه سختی، ناشی از کاتیون‌ها است اما احتمالاً از نظر سختی کربنات (موقت) و غیر کربناته (دائم) نیز مورد بحث قرار می‌گیرد. منابع طبیعی اصلی سختی در آب یون‌های فلزی چند ظرفیتی محلول از سنگ‌های رسوبی، تراوش و رواناب از خاک هستند. در بسیاری از سنگ‌های رسوبی دو یون اصلی کلسیم و منیزیم وجود دارد که متداول ترین آن‌ها سنگ آهک و گچ است. همچنین این ترکیبات از ترکیبات معدنی ضروری مواد غذایی رایج هستند. باتوجه به مطالب بیان شده، سهم کمی در سختی کل آب توسط یون‌های چند ظرفیتی دیگر مانند آلومینیوم، باریم، آهن، منگنز، استرانسیم و روی ایجاد می‌شود [۲].

بسته به pH و قلیایی بودن، سختی بالای ۲۰۰ میلی گرم در لیتر می تواند باعث رسوب، به ویژه در سیستم گرمایشی گردد. آب شیرین با سختی کمتر از ۱۰۰ میلی گرم در لیتر ظرفیت بافر کمتر و امکان خوردگی بیشتر در لوله را دارد. تعدادی از مطالعات اکولوژیکی و اپیدمیولوژی تحلیلی از نظر آماری رابطه مستقیم بین سختی آب آشامیدنی و بیماری های قلبی عروقی را نشان می دهد [۳].

با در نظر گرفتن جمعیت کنونی کشور در ایران، مقدار سرانه آب تجدید شونده حدود ۱۶۷۰ متر مکعب در سال تخمین زده می شود که با این حساب نمی توان آن را یک کشور کم آب تلقی کرد و بهتر است در حال حاضر ایران را جزء کشورهای با تنش آبی در نظر بگیریم. اما توزیع آب در مناطق جغرافیایی کشور نامناسب است. بطوریکه در مقیاس کوچکتر قسمت اعظم مناطق آن کویری، خشک و کم آب می باشد. بخصوص این که جمعیت ایران با نرخ نسبت زیادی در حال رشد است و دیر یا زود ایران نیز در ردیف کشورهای کم آب دنیا قرار خواهد گرفت. با روند کنونی رشد جمعیت و مصرف آب پیش بینی می شود در سال ۱۴۰۰ سرانه آب قابل تجدید کشور به کمتر از ۱۲۵۰ متر مکعب در سال برسد که در این صورت با مشکلات جدی مواجه خواهیم شد [۱].

در مناطق با تراکم جمعیت زیاد، استفاده انسان از منابع آب های زیرزمینی افزایش می یابد که باعث آسیب پذیری منابع می شود. تقریباً هر فعالیتی که به موجب آن احتمال انتشار مواد شیمیایی یا مواد زائد به طور عمدی یا تصادفی در محیط زیست و آلودگی آب های زیرزمینی را به همراه داشته، دارای اهمیت می باشد. هنگامی که آب های زیرزمینی آلوده می شوند، پاک سازی آن بسیار دشوار و گران است. قرار گرفتن در معرض آب سخت از عواملی است که می تواند باعث تشدید آگزاما گردد. عوامل زیادی از جمله گرد و غبار، نایلون، شامپو، تعریق، شنا و لباس باعث تشدید آگزاما می گردد [۴]. با این حال از دلایل عدم قطعیت در مقدار کمی غلظت آب زیرزمینی، می تواند مقدار پارامترهای منطقه مورد مطالعه باشد. بطور کلی برای بررسی محل های آلوده و ارزیابی خطر، شرح دقیق و دستورالعمل مصوب بین المللی بسیار مفید و احتمالاً برای اهداف منطقه ای باید شخصی سازی شود. بنابراین انجام

یک روش مشخص برای تحلیل خطر بگونه‌ای که مطالعات با برنامه ریزی باشد بسیار مفید خواهد بود [۳].

با توجه به موارد ذکر شده، به دلیل کمبود آب‌های سطحی در بسیاری از مناطق و افزایش روز افزون استفاده از منابع آب‌های زیرزمینی؛ مطالعات این منابع بیش از پیش ضروری می‌باشد. یکی از ابزارهای مورد استفاده در مدیریت منابع آب، مدلسازی این منابع می باشد. به عنوان مثال با استفاده از پارامترهای کمی و کیفی می‌توان به پیش‌بینی غلظت آلودگی در آب پرداخت. در صورت وجود داده های مناسب و درک درست از سیستم می‌توان به شیوه‌ای مناسب وضعیت منابع را زیر نظر داشته و در صورت بروز مشکل، در کمترین زمان و مناسب‌ترین هزینه، بهترین روش را برای حل مشکل انتخاب نمود. با وجود مطالعات گسترده بر روی پیش‌بینی غلظت آلاینده‌ها، مطالعات در حوزه سختی آب نسبت به سایر آلاینده‌ها حجم کمتری را به خود اختصاص داده است. در این بین، پیش‌بینی غلظت سختی موقت و دائم آب‌های زیرزمینی با ترکیب شبکه عصبی و سامانه اطلاعات جغرافیایی مورد بررسی قرار خواهد گرفت.

۲-۱ اهداف تحقیق

امروزه یکی از مشکلات اساسی، عدم دسترسی به منابع آبی مناسب برای مصارف گوناگون اعم از شرب، صنعت و کشاورزی است. با توجه به افزایش جمعیت و تقاضای آب برای انواع مصارف به ویژه در کشورهای با اقلیم خشک، نیمه خشک و عدم کیفیت لازم منابع قابل دسترس، باعث انجام تحقیقاتی در مورد کیفیت آب‌های زیرزمینی و میزان کیفیت آن در آینده شده است. بنابراین برای بررسی مشکلات ضروری است یک مدل پیش بینی برای بررسی آن ایجاد گردد.

مدل‌هایی که قادر به آموزش و تعمیم هستند می‌توانند در مسائلی همچون تحلیل، تصمیم‌گیری، تخمین، پیش‌بینی، طراحی و ساخت، مدیریت و کنترل در مطالعات منابع آب مورد استفاده قرار

گیرند. لذا در این تحقیق با داشتن پارامترهای کیفی مانند SO_4^{2-} , Cl^- , PH , TDS , TH , SAR ، EC , K^+ , Na^+ , Mg^{2+} , Ca^{2+} و NO_3^- اقدام به پیش بینی غلظت سختی دائم و موقت در آبخوان دشت آمل و بابل به کمک مدل شبکه عصبی شده است.

بطور عمده هدف از این پژوهش بررسی و شناسایی قابلیت های شبکه عصبی و ساخت مدل بهینه به منظور مدلسازی کیفی و پیش بینی غلظت سختی دائم و موقت در آب های زیرزمینی دشت آمل- بابل می باشد که غلظت سختی دائم و موقت در شبکه عصبی و توزیع مکانی آن توسط سیستم اطلاعات جغرافیایی (GIS) نمایش داده می شود.

۳-۱ روش انجام تحقیق

در این تحقیق به منظور دستیابی به اهداف ذکر شده، ابتدا به جمع آوری اطلاعات و داده های مورد نیاز در مناطق مورد نظر پرداخته، سپس با اندازه گیری پارامترهای کیفی مانند SO_4^{2-} , Cl^- , PH , EC , K^+ , Na^+ , Mg^{2+} , Ca^{2+} و NO_3^- از چاه های آبخوان دشت آمل- بابل در شش سال متوالی جمع آوری گردید. داده های بدست آمده برای ورود به مدل مورد ارزیابی قرار گرفته، سپس با تعیین پارامترهای بهینه برای مدل ANN، مدلسازی غلظت سختی دائم و موقت انجام می شود. جهت مدلسازی، درصدی از داده ها را آموزش داده سپس با نتایج بدست آمده با الگوریتم ماشین بردار پشتیبان مقایسه خواهد شد.

۴-۱ ساختار پایان نامه

در فصل حاضر مقدمه ای بر اهمیت موضوع، اهداف و روش انجام کار مطالعه شد. در فصل دوم پس از معرفی روش های داده محور و توضیح تئوری روش شبکه عصبی، در مورد سایر روش ها نیز توضیح مختصری داده شده است. همچنین در فصل سوم این تحقیق به مرور ادبیات گذشته در این زمینه و

کاربرد شبکه عصبی در مدل سازی آب پرداخته شده است. منطقه مورد مطالعه و داده های مورد استفاده در این تحقیق در فصل چهارم معرفی و پس از ارائه پارامترهای ورودی و خروجی به صورت گام به گام نحوه مدل سازی ارائه و نتایج حاصل بررسی شد. در آخر، در فصل پنجم نتایج و پیشنهادات به منظور ادامه مطالعات در تحقیقات آتی ارائه گردید.

فصل ۲: مبانی و روش تحقیق

۲-۱ مقدمه

مطالعات اولیه در مورد شبکه‌های عصبی مصنوعی نشان دهنده تمایل به داشتن کامپیوتری مشابه یادگیری از انسان است. در نتیجه اصطلاحات مرتبط با ادبیات فنی در این زمینه مملو از عباراتی مانند تحریک و مهار نوروها، استحکام ارتباط سیناپسی، میزان یادگیری، آموزش و تجربه شبکه است. استفاده از شبکه عصبی مصنوعی بعنوان ابزار محاسباتی، مقالات متعددی مرتبط با مزایا و دست آوردهای آن وجود دارد. محققان مدعی هستند علت جذب در شبکه‌های عصبی مصنوعی بدلیل ویژگی‌های مطلوب از تقریب کلی، توانایی یادگیری از نمونه بدون نیاز به ساختار صریح و قابلیت پردازش حجم زیادی از داده‌ها در سرعت‌های بالا می‌باشد. پیش بینی را می‌توان یکی از بیشترین مشکلات هیدرولوژیکی دانست [۵]. بنابراین در بیشتر موارد، استدلال به تنهایی برای تحلیل مسائل

پیچیده مرتبط با کیفیت منابع آب کافی نیست و نیاز به استفاده از روش های مختلف مدلسازی برای پاسخگویی به مقادیر غیر قابل مشاهده می باشد.

بنابراین در بیشتر موارد، استدلال به تنهایی برای تحلیل مسائل پیچیده مرتبط با کیفیت منابع آب کافی نیست و نیاز به استفاده از روش های مختلف مدلسازی برای پاسخگویی به مقادیر غیر قابل مشاهده می باشد.

در این راستا روش های بسیاری برای مدلسازی براساس اطلاعات گذشته وجود دارد. این روش ها شامل شبکه های عصبی مصنوعی، درختان تصمیم گیری، رگرسیون لجستیک، ماشین بردار پشتیبان، رگرسیون چند متغیره می باشد. در ادامه توضیحات بیشتری در خصوص تئوری روش های نام برده و قابلیت های آن ارائه شده است.

۲-۲ مدل های هیدرولوژی

شبکه های عصبی یک مجموعه جایگزین و مکمل مدل های مرسوم هستند. شبکه های عصبی را می توان به عنوان روش های الگوی محاسباتی پویا و تعمیم پذیر تصور کرد که پیش بینی را بدون دانش اولیه از فرایندهای فیزیکی یا شیمیایی، امکان پذیر می کند. از نظر هیدرولوژیست ها این روش به دلیل عدم نیاز به توضیحات دقیق فرآیند از جذابیت قابل توجهی برخوردار است. شبکه های عصبی به ارائه مجموعه داده های کافی متکی هستند و در صورت وجود این شبکه های عصبی می توان برای جستجوی الگوهای موجود در داده ها برنامه ریزی نمود. بر اساس این الگوی تطبیق، ابتدا پیش بینی ها برای مجموعه داده های مستقل برای اعتبارسنجی مدل و سپس برای اهداف عملیاتی انجام می شود. شبکه های عصبی یک رویکرد در چارچوب وسیع تر انفورماتیک آب است که در دهه ۱۹۹۰ به عنوان راهی برای مدیریت اطلاعات با حجم بالا به روشی موثر ظاهر شد. چالش مدیریت آب در ابعاد و کاربردهای فراوان آن تکنیک هایی را می طلبد که بتوانند اجزای بی شماری از پیچیدگی های هیدرولیک و کیفیت آب تا برنامه های مالی و برنامه های اجتماعی را به هم پیوند دهند.

تکنیک‌های موجود برای هیدرولوژیست‌ها بسیار زیاد و متنوع است که هرکدام مزایا و معایب خاص خود را دارند. از چالش‌های پیش‌بینی می‌توان به تنوع محیط طبیعی، حوضه‌های آبریز متفاوت، نامنظمی سیستم و پیچیدگی ادغام مقیاس به همراه هزینه به دست آوردن داده‌ها اشاره نمود [۵].

بیشتر فرآیندهای هیدرولوژیکی غیرخطی هستند و درجه بالایی از تنوع مکانی و زمانی را نشان می‌دهند.

به‌طور کلی مدل‌های هیدرولوژیکی که امروزه مورد استفاده قرار می‌گیرند به سه دسته مدل‌های تجربی، ژئومورفولوژیکی و مدل‌های فیزیکی تقسیم بندی می‌شوند:

- مدل‌های تجربی، سیستم‌های هیدرولوژیکی را به عنوان یک جعبه سیاه^۱ در نظر گرفته و سعی در ایجاد رابطه بین تاریخچه ورودی و خروجی هستند. این روش‌ها به تاریخچه طولانی احتیاج دارند و هیچگونه مبنای فیزیکی ندارند؛ به همین ترتیب برای حوضه‌های آبریز قابل استفاده هستند.

- مدل‌های ژئومورفولوژی ساختار حوضه و شبکه جریان را به خوبی نشان می‌دهند، اما لازم است فرضیه‌های مختلفی در مورد خطی بودن پاسخ واحدهای حوزه آبخیزداری ارائه شود.

- به‌طور کلی مدل‌های فیزیکی شامل حل یک سیستم از معادلات دیفرانسیل با مشتقات جزئی هستند که درک درستی از فرآیند جریان حوزه را به ما نشان می‌دهند. در بیشتر مشکلات، یک حل عددی مطلوب با گسسته سازی ابعاد مکان-زمان از گره‌های مجزا جستجو می‌شود. این نوع داده‌ها حتی در حوزه‌های تحقیقاتی با ابزار دقیق به ندرت در دسترس هستند. مشکل تنوع مکانی و نمایش دقیق حوضه آبخیز منجر به ارائه مدل‌های حوضه تصادفی برای پاسخگویی به این عدم قطعیت شده است. از این مدل‌ها در مطالعات برای نشان دادن پاسخ غیرخطی استفاده شد. این مدل‌ها مشکلاتی مانند شناسایی، تخمین پذیری و متفاوت بودن چگونگی تخمین پارامتر را دارا می‌باشند. حتی با وجود توانایی‌های محاسباتی فعلی نمایش فیزیکی حوزه آبخیز در مدل، در بهترین حالت، یک تقریب است.

با وجود این محدودیت‌ها، ثابت شده است که این مدل‌ها در صورت استفاده مناسب برای بسیاری از مشکلات هیدرولوژیکی بسیار مفید هستند [۶].

موارد زیر برخی از دلایلی است که شبکه عصبی را به یک ابزار محاسباتی جذاب تبدیل کرده است:

- شبکه عصبی قادرند رابطه بین متغیرهای ورودی و خروجی را بدون ملاحظه فیزیکی صریح تشخیص دهند.

- در صورتی که اندازه‌گیری‌ها دارای خطا و اختلال باشد شبکه عصبی به خوبی عمل می‌کند.
- شبکه عصبی قادر به انطباق با راه‌حل‌های سازگار در گذر زمان بر اثر تغییر شرایط می‌باشد.
- شبکه عصبی به‌طور ذاتی دارای خصوصیات پردازش اطلاعات هستند و پس از آموزش، استفاده از آن‌ها آسان است.

اگرچه، مطالعات انجام شده زیادی، توانایی شبکه عصبی را در این نوع مسائل نشان می‌دهد، اما محدودیت‌های این روش را نباید نادیده گرفته شود [۶].

از محدودیت‌های این روش می‌توان به وابستگی نتایج به داده‌های پیشین اشاره نمود. علت این محدودیت، عدم ثبت اطلاعات کافی از داده‌ها و یا عدم اطمینان از یکسان بودن شرایط در صورت وجود داده، در بسیاری از مسائل می‌باشد. یکی دیگر از محدودیت‌های مهم شبکه عصبی، فقدان مفاهیم و روابط فیزیکی است که یکی از دلایل اصلی آن نگرش بدبینانه نسبت به این روش می‌باشد. این واقعیت که هیچ روش استاندارد برای انتخاب معماری شبکه وجود ندارد نیز انتقاداتی را به همراه دارد [۶].

در ادامه به توضیح کامل تئوری روش شبکه عصبی و ماشین بردار پشتیبان پرداخته، سپس کلیاتی در مورد سختی آب توضیح داده شده است.

۲-۳ شبکه‌های عصبی مصنوعی (ANN)

بسیاری از وظایف مربوط به تشخیص فهم یا الگو به صورت خودکار بسیار دشوار است، اما به نظر می‌رسد توسط انسان بسیار آسان انجام می‌شود. برای مثال با کمترین تلاش، اشیاء مختلف را تشخیص و تحت تاثیر اطلاعات بصری محیط پیرامون خود قرار می‌گیرد.

این شبکه به عنوان تعمیم مدل‌های ریاضی شناخت انسان یا زیست شناسی عصبی براساس فرضیات زیر می‌باشد:

-پردازش اطلاعات در بسیاری از عناصر ساده به نام نورون اتفاق می افتد.

-هر اتصال دارای یک وزن مرتبط است که در یک شبکه عصبی معمولی سیگنال منتقل شده را چند برابر می‌کند.

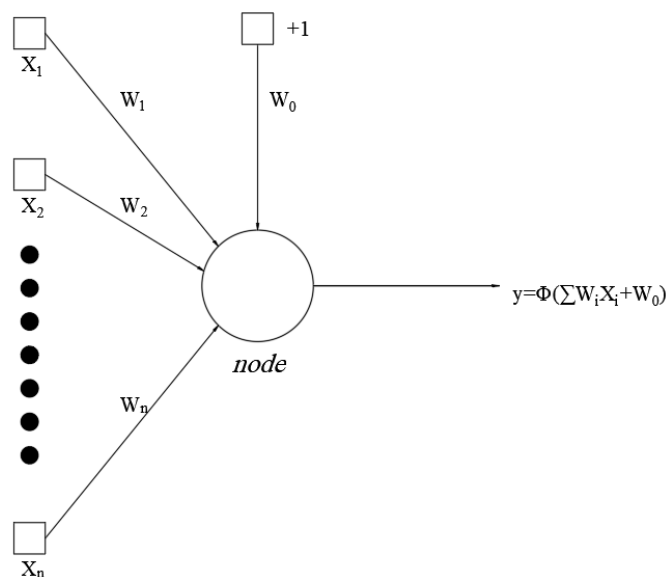
-هر نورون تابع فعال به ورودی شبکه خود، برای تعیین سیگنال خروجی آن اعمال می‌شود.

-سیگنال‌ها از طریق اتصال بین نورون‌ها منتقل می‌شوند.

یک شبکه عصبی شامل تعداد زیادی از عناصر پردازشی ساده به نام نورون، سلول یا گره است. هر نورون با استفاده از پیوندهای ارتباطی هدایت شده هر یک با یک وزن مرتبط به سلول‌های عصبی دیگر متصل می‌شود. نورون حالتی دارد که به آن لایه میانی می‌گویند این لایه تابعی از ورودی‌ها می‌باشد. به طور معمول، یک نورون فعالیت خود را به عنوان سیگنال به چندین نورون دیگر ارسال می‌کند. اگرچه این سیگنال همزمان به چندین نورون دیگر نیز منتشر می‌شود اما نکته مهم این است که یک نورون می‌تواند فقط یک سیگنال ارسال کند [۷].

۱-۳-۲ ساختار شبکه عصبی

در شبکه‌های عصبی، مدل‌های محاسباتی یا گره‌ها از طریق وزن‌ها برای سازگاری در بهبود عملکرد در حین استفاده متصل می‌شوند. ایده اصلی، دستیابی به عملکرد مناسب از طریق تراکم اتصال داخلی عناصر محاسباتی ساده می‌باشد. ساده‌ترین گره ترکیبی از N وزن $W_1, W_2, \dots, W_n$ و N ورودی $X_1, X_2, \dots, X_n$ و نتیجه را از یک Φ غیر خطی فراهم می‌کند. از نظر الگوی اتصال، شبکه‌های عصبی را می‌توان به دو دسته تقسیم کرد: شبکه‌های عصبی پیشرو، یک شبکه عصبی مصنوعی است که در آن اتصال میان واحدهای تشکیل دهنده، یک چرخه را تشکیل نمی‌دهند. در واقع این شبکه متفاوت از شبکه‌های عصبی بازگشتی می‌باشد. شبکه عصبی پیشرو اولین و ساده‌ترین نوع شبکه عصبی مصنوعی می‌باشد. در این شبکه اطلاعات تنها از یک مسیر حرکت می‌کند که جهت آن رو به جلو می‌باشد. در واقع اطلاعات با شروع از گره (نورون)‌های ورودی و گذر از لایه‌های پنهان (در صورت وجود) به سمت گره‌های خروجی می‌روند. شبکه عصبی بازگشتی که در آن اتصالات بین گره‌هایی از یک گراف جهت‌دار در امتداد یک سری زمانی می‌باشند و سبب می‌شود تا الگوریتم بتواند بطور موقت رفتار پویا را به نمایش بگذارد [۸].



شکل ۱-۲. اتصالات نورون [۸]

مطابق شکل ۲-۳ معادله

$$Y = \Phi \left(\sum_{i=1}^N w_i x_i + w_0 \right) \quad (۱-۲)$$

نشان دهنده شکل ریاضی یک نورون می‌باشد. بردار ورودی توسط $X = [x_1, \dots, x_N, 1]^T$ و $w = [w_1, \dots, w_N, w_0]^T$ به عنوان بردار وزن نورون نامیده می‌شود. وزن w_0 وزنی است که برابر با ورودی بایاس بوده که به طور معمول روی واحد تنظیم می‌شود. تابع ۱) $\Phi: \mathbb{R} \rightarrow \mathbb{R}$ یکنواخت و پیوسته می‌باشد و معمولاً به شکل سیگنویید نمایش داده می‌شود. [۸]

۲-۳-۲ آموزش شبکه

به منظور انجام یک شبکه عصبی مصنوعی برای ایجاد بردار $Y = (y_1, y_2, \dots, y_p)$ که تا حد ممکن به بردار هدف $T = (t_1, t_2, \dots, t_p)$ نزدیک باشد نیاز به فرآیندی به نام آموزش است که یادگیری نیز نامیده می‌شود. این فرایند برای پیدا کردن ماتریس‌های وزنی بهینه و بردارهای بایاس، که یک تابع خطای از پیش تعیین شده را که معمولاً به شکل زیر است به حداقل می‌رساند.

$$E = \sum_p \sum_p (y_i - t_i)^2 \quad (۲-۲)$$

در رابطه بالا t_i ، یک مولفه از خروجی مورد نظر T ، y_i برابر خروجی مربوط به شبکه عصبی مصنوعی و P تعداد الگوهای آموزش می‌باشد. آموزش، فرایندی است که به وسیله آن وزن اتصال شبکه عصبی مصنوعی از طریق یک روند مداوم موثر توسط محیطی که شبکه در آن تعبیه شده است، سازگار می‌شود. اصولاً دو نوع آموزش با نظارت و بدون نظارت وجود دارد. یک الگوریتم آموزش با نظارت به یک آموزش دهنده خارجی برای راهنمایی فرایند آموزش نیاز دارد. هدف اصلی آموزش به حداقل رساندن تابع خطا با جستجو برای مجموعه‌ای از نقاط قوت اتصال و مقادیر قابل تحمل است که باعث ایجاد خروجی‌هایی نزدیک به هدف در شبکه عصبی می‌شود. پس از اتمام آموزش امید است

شبکه عصبی باتوجه به ورودی‌های جدید بتواند نتایج منطقی ایجاد کند. در طرف مقابل یک الگوریتم یادگیری بدون نظارت وجود دارد که نیاز به آموزش دهنده ندارد. در طول آموزش فقط یک مجموعه داده ورودی به شبکه عصبی ارائه می‌شود که به طور خودکار وزن اتصال آن را تنظیم می‌کند تا الگوها را در کلاس‌هایی با مشخصات مشابه دسته بندی کند [۶].

۳-۲-۳ روش‌های یادگیری

یادگیری یکی از توانایی مهم در شبکه عصبی است. قواعد یادگیری الگوریتم برای یافتن وزن مناسب W و یا سایر پارامترهای شبکه هستند. یادگیری یک شبکه عصبی را می‌توان به عنوان یک مسئله بهینه سازی غیرخطی مشاهده کرد که هدف از آن یافتن مجموعه‌ای از پارامترهای شبکه برای به حداقل رساندن تابع هزینه برای مثال‌های ارائه شده می‌باشد. به عبارت دیگر، یادگیری یک فرایند بهینه سازی است که با تنظیم پارامترهای شبکه مانند W ، منجر به ایجاد خروجی می‌شود که تا حد ممکن به خروجی مورد نظر نزدیک باشد. به این نوع تخمین، پارامتر الگوریتم یادگیری یا آموزش نیز می‌گویند. شبکه‌های عصبی معمولاً بصورت یک دوره آموزش می‌بینند. یک دوره^۱ به یک تکثیر رفت و برگشت در شبکه اشاره دارد. یعنی یک دوره برابر با یک رفت و برگشت ورودی در کل شبکه است.

روش‌های یادگیری به طور متعارف به یادگیری نظارت شده، بدون نظارت، تقویتی^۲ و فرگشتی^۳ تقسیم می‌شوند. به طور گسترده یادگیری با نظارت در تشخیص الگو، تقریب، کنترل، مدلسازی و شناسایی، پردازش سیگنال و بهینه‌سازی و یادگیری بدون نظارت عمدتاً برای شناسایی الگو، خوشه‌بندی، کمی سازی بردار، کدگذاری سیگنال و تجزیه و تحلیل داده‌ها استفاده می‌شود. پس از یادگیری، شبکه عصبی نمایانگر یک رابطه پیچیده است و توانایی تعمیم را دارد. زمانی که ورودی جدیدی به شبکه عصبی تربیت شده داده می‌شود، یک خروجی منطقی را تولید می‌کند [۷].

^۱Epoch

^۲Reinforcement

^۳Evolutionary

۴-۳-۲ یادگیری نظارت شده

یادگیری نظارت شده بر مبنای مقایسه بین خروجی واقعی شبکه و خروجی مورد نظر است که پارامترهای آن با ترکیبی از مجموعه الگوی آموزش و خطاهای مربوطه بین خروجی مورد نظر و پاسخ واقعی شبکه تنظیم می‌شوند. برای یادگیری نظارت شده، از یک معیار خطا، که تفاوت بین خروجی شبکه و خروجی از نمونه های آموزش را نشان می‌دهد، برای هدایت فرایند یادگیری استفاده می‌شود. اندازه گیری خطا معمولاً با میانگین مربع خطای تعریف می‌شود.

$$E = \frac{1}{N} \sum_{p=1}^N \|y_p - \hat{y}_p\|^2 \quad (3-2)$$

در رابطه بالا، N ، برابر تعداد زوج الگو در S مجموعه نمونه، y_p خروجی قسمتی از P امین زوج الگو و $\hat{y}_p$ خروجی شبکه مربوط به جفت الگو P می‌باشد. فرایند یادگیری هنگامی خاتمه می‌یابد که E به کوچک‌ترین اندازه خود برسد. همچنین برای کاهش E ، معمولاً از روش گرادینان نزولی^۱ استفاده می‌شود [۷].

در حالت یادگیری با نظارت سیستم برای هر نمونه‌ی آموزشی یادگیری را براساس یک المان دوتایی از ورودی و برچسب که خروجی مساله می‌باشد، فرا می‌گیرد. مجموعه داده دارای برچسب به معنی این است که برای هر عضوی از مجموعه آموزشی، یک پاسخ یا راه حل نیز برای آن داده شده است. هدف در اینجا، تطبیق سیستم، به گونه‌ای است که برای ورودی جدید سیستم بتواند خروجی درست را براساس آنچه تاکنون از داده‌های آموزشی فراگرفته است، پیش‌بینی کند. در یادگیری نظارتی اگر داده‌های مسئله جهت یادگیری به صورت گسسته باشند، این مساله دسته بندی و اگر مقادیر داده ها به صورت پیوسته باشند به آن رگرسیون گویند.

^۱MSE

^۲Gradient-Descent

۵-۳-۲ یادگیری بدون نظارت

یادگیری بدون نظارت نوعی یادگیری ماشین است که شامل هیچ هدفی نمی‌شود. در حقیقت تلاش می‌کند تا اطلاعات ورودی را به طور خودکار، بدون خروجی مشخص با کاهش حقیقی ابعاد داده یا مقدار کل داده‌های ورودی مرتبط کند. یادگیری بدون نظارت فقط براساس همبستگی بین داده‌های ورودی است و برای یافتن الگوها یا ویژگی‌های قابل توجه در داده‌های ورودی بدون کمک راهنما^۱ استفاده می‌شود [۷]. خوشه بندی نمونه‌ای از این نوع یادگیری می‌باشد.

۶-۳-۲ پرسپترون چندلایه (MLP)

پرسپترون چند لایه به یک طبقه کلی از ساختار به نام شبکه های عصبی پیشرو تعلق دارند، یک نوع اساسی از شبکه عصبی که قادر به تقریب طبقات عمومی توابع از جمله توابع پیوسته و یکپارچه است [۹].

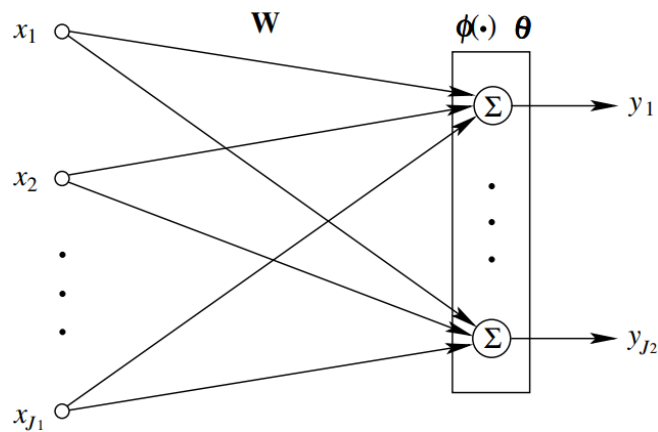
در الگوریتم پرسپترون، هنگام آموزش برای هر بردار ورودی، خروجی شبکه را تولید کرده و آن را با مقدار درست مقایسه می‌کند تا مقدار خطا مشخص شود. در صورتی که در این شبکه خطایی رخ ندهد وزن‌ها ثابت و آموزش تا زمانی که شبکه بدون خطا شود ادامه پیدا خواهد کرد.

۱-۶-۳-۲ پرسپترون تک لایه

برای پرسپترون تک لایه، توپولوژی شبکه همانند شکل ۲-۲ بوده و شبکه ورودی به نورون توسط رابطه

$$net = \sum_{i=1}^{J_1} w_i x_i - \theta = w^T x - \theta \quad (۴-۲)$$

نشان داده می‌شود.



شکل ۲-۲. معماری پرسپترون تک لایه [۷]

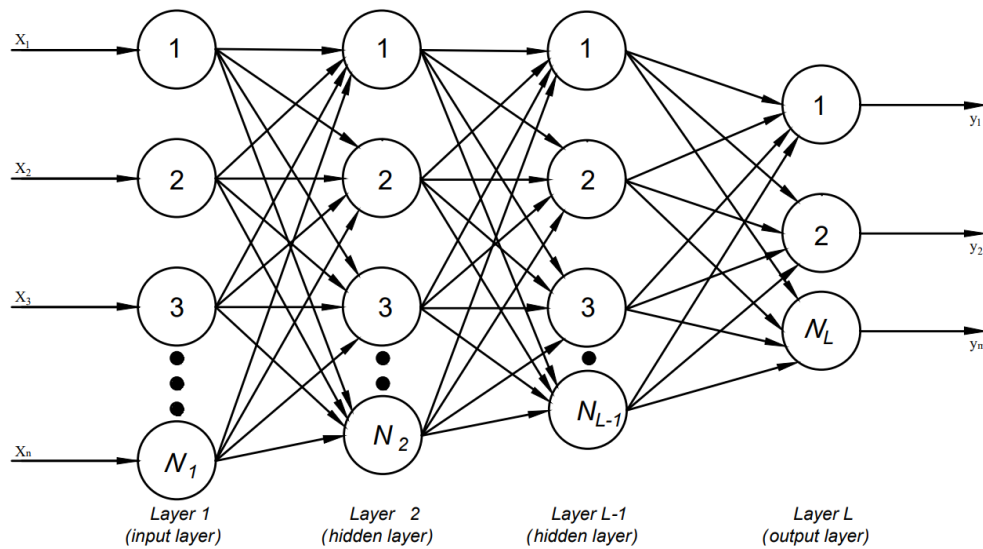
پرسپترون تک لایه با استفاده از تابع فعالسازی Hard-Limiter برای طبقه‌بندی بردار x در دو دسته، قابل استفاده است. دو منطقه تصمیم‌گیری توسط یک صفحه از یکدیگر جدا شده‌اند. که رابطه آن به صورت زیر نشان داده می‌شود.

$$w^T x - \theta = 0 \quad (۵-۲)$$

در رابطه بالا θ پارامتری است که برای تغییر مرز تصمیم‌گیری از مبدا استفاده می‌گردد [۷].

۲-۳-۶-۲ ساختار پرسپترون چند لایه

در ساختمان پرسپترون چند لایه، نورون‌ها در هر لایه دسته بندی می‌شوند. لایه‌های ابتدایی و انتهایی را به ترتیب لایه ورودی و خروجی گویند. به لایه‌های باقیمانده، لایه‌های پنهان گفته میشود. به طور معمول یک شبکه عصبی پرسپترون چند لایه از یک لایه ورودی، یک یا چند لایه پنهان و یک خروجی تشکیل شده که در شکل ۳-۲ نشان داده شده است [۹].



شکل ۳-۲. ساختار پرسپترون چند لایه [۹]

با فرض تعداد لایه‌های کل برابر L ، اولین لایه، لایه ورودی، L امین لایه، لایه خروجی و لایه ۲ تا $L-1$ ، لایه مخفی نام دارد. بنابراین w_{ij}^l نشان دهنده وزن پیوند بین j امین نورون از $L-1$ امین لایه و i امین نورون از L امین لایه که $1 \leq i \leq N_l$ ، $1 \leq j \leq N_{l-1}$ می‌باشد. همچنین x_i نشان دهنده i امین ورودی خارجی به MLP و z_i^l خروجی i امین نورون از L امین لایه می‌باشد. یک پارامتر اضافی وزن برای هر نورون w_{i0}^l معرفی می‌کنیم که نشان دهنده تاثیر آن بر روی نورون i ام از لایه L است. بنابراین w از پرسپترون چند لایه شامل w_{ij}^l ، $j=0,1,\dots,N_{l-1}$ و $i=1,2,\dots,N_l$ و $l=1,2,3,\dots,L$ ، به صورت زیر است: [۹]

$$\mathbf{w} = [w_{10}^2, w_{11}^2, w_{12}^2, \dots, w_{N_L N_{L-1}}^L]^T \quad (۶-۲)$$

۷-۳-۲ شبکه عصبی پیشرو^۱

باتوجه به استفاده از شبکه عصبی پیشرو در این مطالعه، لازم است تابع یادگیری و آموزش مورد بررسی قرار گیرند. در تحقیقات زیادی کارایی انواع تابع یادگیری و آموزش مورد مطالعه قرار گرفته است. از میان توابع موجود برای یادگیری شبکه، گرادیان کاهشی با وزن مومنتوم عملکرد مناسبی را نسبت به سایر توابع یادگیری دارد [۱۰]. بعلاوه از تابع آموزش لوببرگ-مارکوارت^۲ که ترکیبی از دو الگوریتم روش گرادیان کاهشی و روش گاوس-نیوتون است، برای شبکه عصبی به عنوان الگوریتمی کارآمد در نظر گرفته شد [۱۱].

شکل ۲-۴ یک شبکه عصبی پیشرو با سه لایه را نشان می‌دهد. تعداد المان های پردازش در لایه‌های ورودی و خروجی بستگی به نوع مسئله دارد، درحالی که برای لایه پنهان تعداد آن بر اساس آزمون و خطا تعیین می‌گردد. لایه ورودی، متغیر ورودی را دریافت و اطلاعات ضروری را به المان‌های پردازش لایه اول ارائه می‌دهد. گره‌های موجود در یک لایه فقط به قسمت‌های بعدی متصل و جهت جریان اطلاعات در شبکه فقط از لایه ورودی به لایه خروجی می باشد. مقادیر خروجی لایه پنهان المان پردازش به صورت معادله زیر ارائه می‌گردد:

$$y'_j = g \left[\sum_{i=1}^{n_{in}} x_i w_{ij} + b_j \right]; j = 1, \dots, n_{hid} \quad (۷-۲)$$

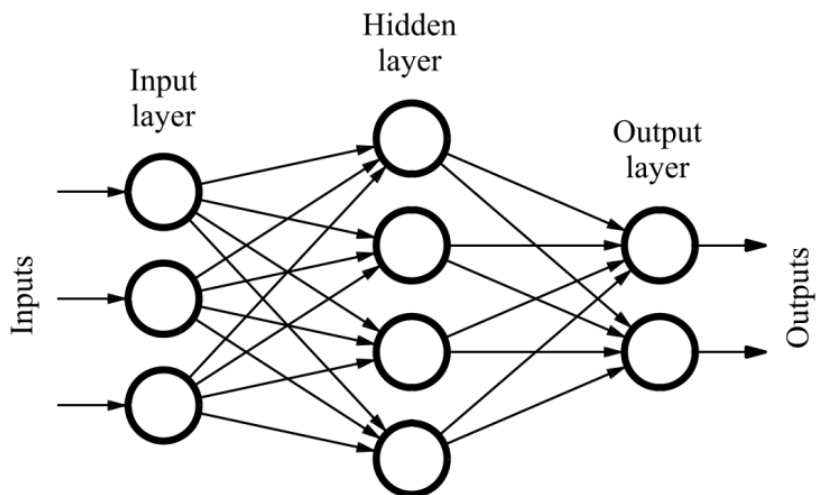
در این رابطه، n_{in} و n_{hid} به ترتیب تعداد لایه ورودی و لایه پنهان، w_{ij} وزن اتصال مرتبط به i امین المان پردازش در لایه ورودی و j امین المان پردازش در لایه پنهان، b_j برابر بایاس برای j امین المان پردازش در لایه پنهان، x_i برابر i امین مولفه از بردار ورودی n بعدی و $g(.)$ تابع انتقال المان پردازش پنهان می‌باشد. خروجی شبکه مورد نظر به صورت زیر تعریف می‌گردد:

$$y_k = f \left[\sum_{i=1}^{n_{hid}} y'_i w_{ik} + b_k \right]; j = 1, \dots, n_{out} \quad (۸-۲)$$

^۱FeedForward Neural Network

^۲Levenberg Marquardt

در این رابطه n_{out} برابر تعداد المان‌های پردازش در لایه خروجی، w_{jk} برابر وزن اتصال مرتبط با j امین المان پردازش در لایه پنهان و k امین المان پردازش در لایه خروجی، b_k ، بایاس برای k امین المان عنصر در لایه خروجی و $f(.)$ تابع انتقال k امین خروجی المان پردازش می‌باشد. تابع تبدیل تعیین کننده پاسخ گره به ورودی دریافتی خالص می‌باشد [۱۲].



شکل ۲-۴. نمونه‌ای از شبکه عصبی پیشرو [۱۳]

۸-۳-۲ الگوریتم پرسرو

الگوریتم یادگیری شبکه عصبی پرسرو، تعمیم یافته قانون دلتا است که به الگوریتم حداقل میانگین مربعات^۱ نیز معروف است. پرسرو بر محدودیت‌های یادگیری پرسپترون که توسط مینسکی و پاپرت^۲ برشمرده شده است غلبه می‌کند. باتوجه به الگوریتم پرسرو، MLP توانایی تعمیم به لایه‌های با تعداد بالا را دارد. الگوریتم پرسرو خطای بین سیگنال مورد نظر و خروجی شبکه را از طریق شبکه به عقب منتقل می‌کند. پس از ارائه الگوی ورودی، سپس خروجی شبکه با یک الگوی هدف داده شده مقایسه و خطای خروجی هر واحد محاسبه می‌شود. این سیگنال خطا را به عقب انتشار داده و در نهایت یک سیستم کنترل حلقه بسته^۳ ایجاد می‌شود. الگوریتم پرسرو یک روش گرادیان نزولی نظارت شده است که در آن MSE بین خروجی شبکه و خروجی واقعی مورد نظر به حداقل می‌رسد [۷].

هدف اصلی در ایجاد مدل عصبی، یافتن مجموعه‌ای بهینه از پارامترهای w ، به گونه‌ای که $y = y(x, w)$ ، رفتار اصلی مسئله را با تخمینی نزدیک نشان می‌دهد. این امر از طریق فرایندی به نام آموزش حاصل می‌شود (یعنی بهینه‌سازی در فضای w). مجموعه‌ای از داده‌ها برای آموزش به شبکه عصبی ارائه می‌شود. در طول آموزش، عملکرد شبکه عصبی اختلاف بین خروجی واقعی شبکه عصبی و خروجی مطلوب برای تمام نمونه‌های آموزش مورد محاسبه قرار می‌گیرد که تفاضل آن به عنوان خطا شناخته می‌شود [۹].

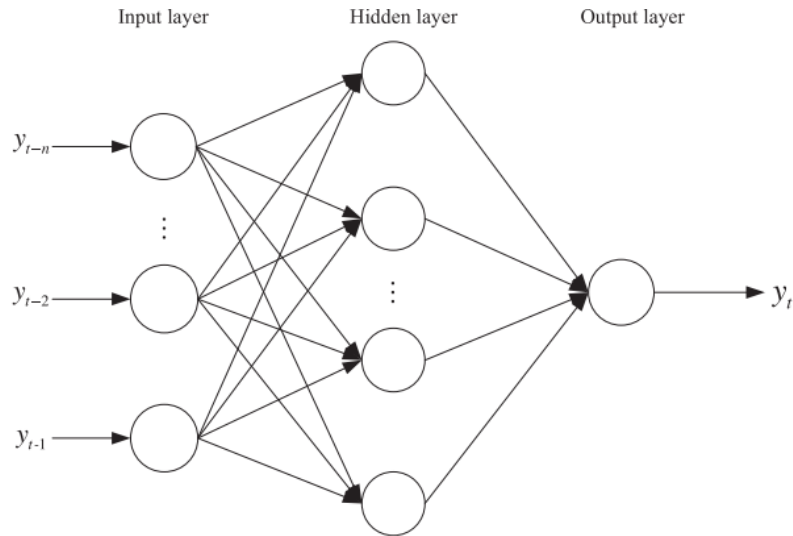
$$E = \frac{1}{2} \sum_{k \in T_r} \sum_{j=1}^m (y_j(x_k, w) - d_{jk})^2 \quad (9-2)$$

از رابطه بالا d_{jk} ، j امین المان از d_k ، j امین خروجی شبکه عصبی برای ورودی x_k و T_r یک شاخص از داده‌های آموزش می‌باشد. پارامتر وزنی w در طول آموزش تنظیم و خطای آن به حداقل می‌رسد.

^۱Least Mean Squares

^۲Minsky and Papert

^۳Closed Loop Control



شکل ۲-۵. نمونه‌ای از شبکه عصبی پسر [۱۴]

۲-۴ تابع فعال‌ساز

شبکه‌های عصبی، شبکه‌ای از چند لایه سلول عصبی متشکل از گره‌هایی است که برای طبقه‌بندی و پیش‌بینی برخی داده‌ها به عنوان ورودی به شبکه استفاده می‌شود. تابع فعال‌ساز در شبکه عصبی از یک خروجی که به سادگی تابع خطی است جلوگیری می‌کند. اگرچه حل یک معادله خطی ساده و آسان است اما پیچیدگی آن‌ها محدود است و آن‌ها توانایی یادگیری و تشخیص نگاشت‌های پیچیده را از داده‌ها ندارند [۱۵].

معمولاً همه نورون‌ها در MLP از همان تابع فعال‌سازی سیگموئیدال استفاده می‌کنند. این عمل باعث ایجاد محدودیت اعداد در بازه $(0, 1)$ یا $(-1, 1)$ می‌شود. با این حال، با توجه مشکلات تقریب تابع ممکن است خروجی شبکه از خروجی مورد نظر فاصله داشته باشد؛ در واقع الگوریتم آموزش به درستی عمل نکرده و نامعتبر است [۷].

۱-۴-۲ تابع Linear

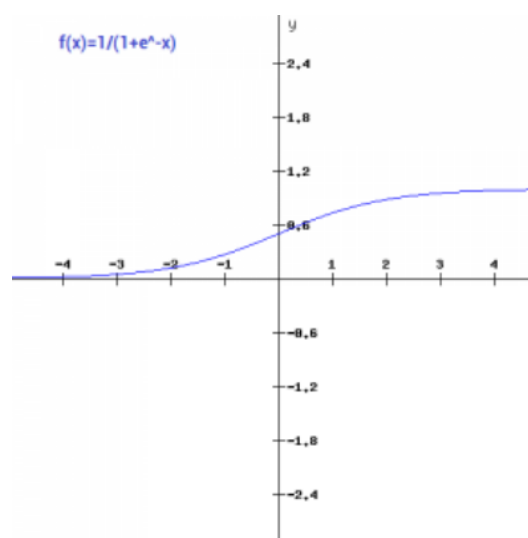
در صورت انتخاب تابع خطی $\varphi(x) = x$ در لایه خروجی برای نورون‌ها به عنوان تابع فعال‌سازی برای افزایش دامنه دینامیکی خروجی شبکه، مراحل پیش‌پردازش و پس‌پردازش لازم نیست. زمانی که تابع فعال‌سازی لایه خروجی به عنوان $\varphi(x) = x$ انتخاب شود، رابطه بین آخرین لایه پنهان و لایه خروجی خطی است، بنابراین شبکه را می‌توان در دو مرحله آموزش داد [۷].

۲-۴-۲ تابع Sigmoid

تابع سیگموئید به دلیل غیرخطی بودن به طور گسترده مورد استفاده قرار می‌گیرد. این تابع مقادیر را بین صفر تا یک تبدیل می‌کند. رابطه آن به صورت زیر تعریف می‌شود.

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad (۱۰-۲)$$

تابع سیگموئید در مورد صفر برابر نیست، این بدین معنی است که اثر تمام مقادیر خروجی نورون‌ها یکسان خواهند بود. این مسئله را با مقیاس گذاری تابع سیگموئید می‌توان بهبود بخشید [۱۵].



شکل ۲-۶. تابع سیگموئید [۱۵]

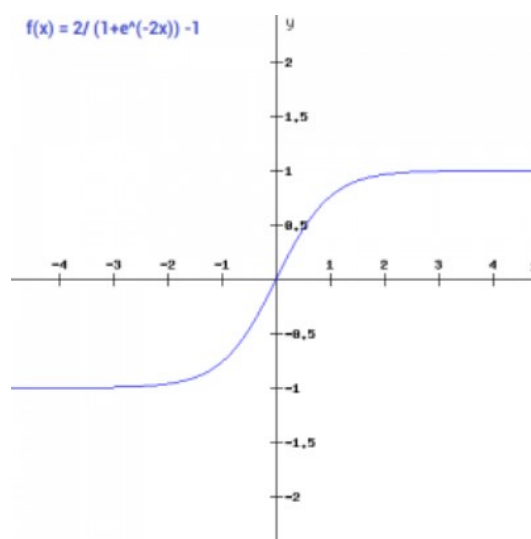
۳-۴-۲ تابع Tanh

این تابع، یک تابع هیپربولیک است که مشابه تابع سیگموئید بوده اما تفاوت آن تقارن این تابع نسبت به مبدا می باشد.

ورودی لایه بعدی از خروجی لایه قبلی تغذیه می کند که منجر به علائم مختلف خروجی از لایه های قبلی می شود. این تابع را می توان به صورت زیر تعریف نمود:

$$\sigma(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (۱۱-۲)$$

در این تابع مقادیر بین یک و منفی یک قرار دارد. همچنین در مقایسه با تابع سیگموئید، گرادیان تابع Tanh دارای شیب بیشتری می باشد [۱۵].



شکل ۲-۷. تابع تانژانت هیپربولیک [۱۵]

۵-۲ ماشین بردار پشتیبان

ماشین بردار پشتیبان یک شبکه عصبی سه لایه است که فرمول بندی آن بر اساس اصل کمینه سازی ساختاری در تئوری یادگیری آماری می باشد. [۷] این روش نسبت به اصل کمینه سازی خطای تجربی که معمولاً مورد استفاده قرار می گیرد، برتری دارد. [۱۶] کمینه سازی خطای ساختاری کران بالای خطا را کمینه کرده در حالی که کمینه سازی خطای تجربی به کمینه سازی خطای داده های آموزشی می پردازد [۱۶].

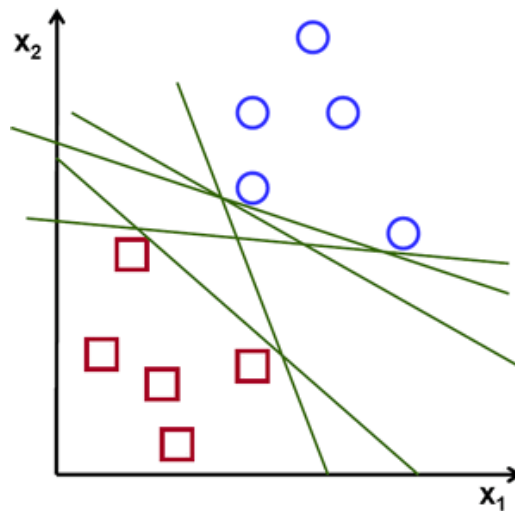
استفاده از SVM به عنوان رگرسیون در سال ۱۹۹۵ توسط وپنیک مطرح شد. [۱۷] در روش رگرسیون، جستجو برای یافتن تابع هدف به صورتی است که همزمان با بیشینه نمودن کارایی و خاصیت تعمیم پذیری مدل، مقدار پیچیدگی آن کمینه گردد. این عمل با بهره گیری از داده های آموزشی ورودی و خروجی بگونه ای که مجموعه ای از M جفت داده های ورودی و خروجی از $i=1,2,\dots,M$ به سیستم داده شده و سیستم تلاش می کند تا تابع را از ورودی به خروجی به گونه ای آموزش دهد که توان تخمین مقادیر مشاهده نشده x را برای خروجی های y داشته باشد. به عبارت دیگر باتوجه به روش ارائه شده، می خواهیم مجموعه توابع وابسته به x را به نحوی بیابیم که در مجموع، کمترین میزان انحراف را به ازای هر داده داشته بطوری که تابع تا حد امکان بصورت هموار باشد [۱۸].

باتوجه به نوع مسئله، در این تحقیق از فرم رگرسیونی ماشین بردار پشتیبان استفاده شده است. در ادامه به توضیحاتی از این روش خواهیم پرداخت.

۲-۶ دسته بندی بردار پشتیبان

مسئله دسته بندی می‌تواند یک مسئله دو کلاسه، بدون آنکه کلیات خطای آن تغییر نماید، باشد. در اینگونه مسائل هدف از جداسازی دو کلاسه با استفاده از یک تابع، براساس مثال‌های موجود، تعیین می‌گردد. هدف از تولید دسته بندی، کاربرد آن‌ها در مثال‌های غیر نظارت شده می‌باشد. در اینجا طبقه بندی‌های خطی زیادی وجود دارد که توانایی جداسازی داده‌ها را دارند، اما تنها یکی از آن‌ها وجود دارد که حاشیه را به حداکثر می‌رساند (حداکثر فاصله بین جدا کننده و نزدیکترین نقطه داده در هر کلاس). به این طبقه بندی خطی، ابر صفحه جدا کننده بهینه گفته می‌شود [۱۶].

شکل ۲-۸ نمونه‌ای از دسته بندی خطی است که توسط ابر صفحه جدا شده‌اند.



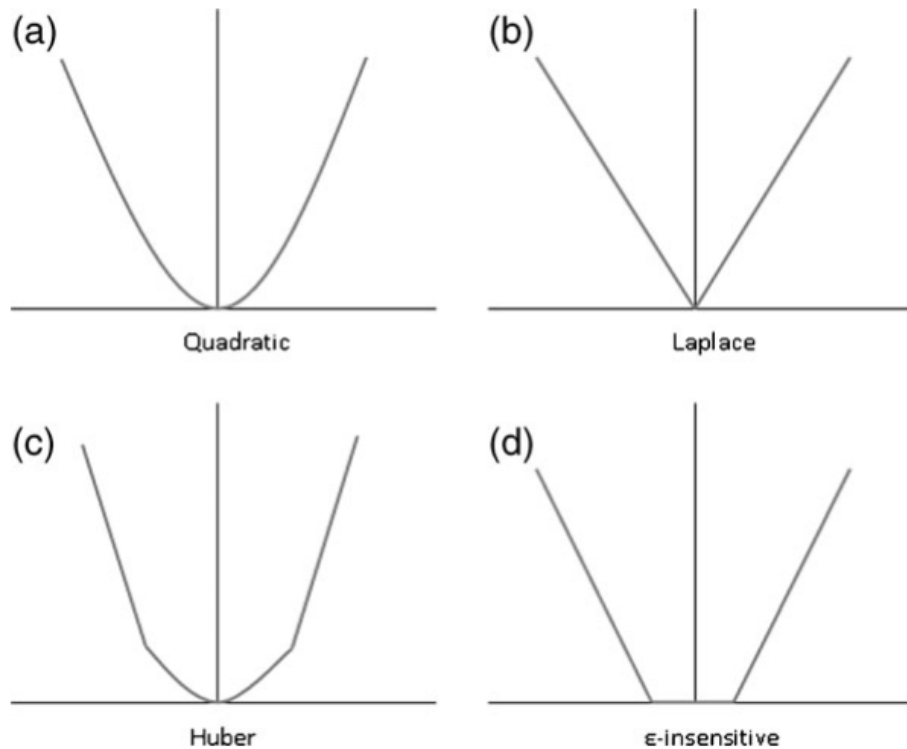
شکل ۲-۸. ابر صفحه جدا کننده بهینه [۱۹]

۲-۷ رگرسیون بردار پشتیبان

رگرسیون بردار پشتیبان یک شاخه مهم از ماشین بردار پشتیبان است که تفاوت آن با دسته بندی ماشین بردار پشتیبان تنها در یک نقاط نمونه می‌باشد. در این روش ابر صفحه بهینه به دنبال به حداکثر رساندن فاصله بین دو یا چند نوع از تعداد نقاط مانند SVM نیست، بلکه به دنبال به حداکثر رساندن مجموع انحراف بین نقاط نمونه یا ابر صفحه می‌باشد [۲۰].

۲-۷-۱ حالت خطی رگرسیون بردار پشتیبان

ماشین بردار پشتیبان توانایی حل مسائل رگرسیونی را با ایجاد تابع خطا^۱ دارد. شکل ۹-۲ چهار تابع خطا را نشان می‌دهد.

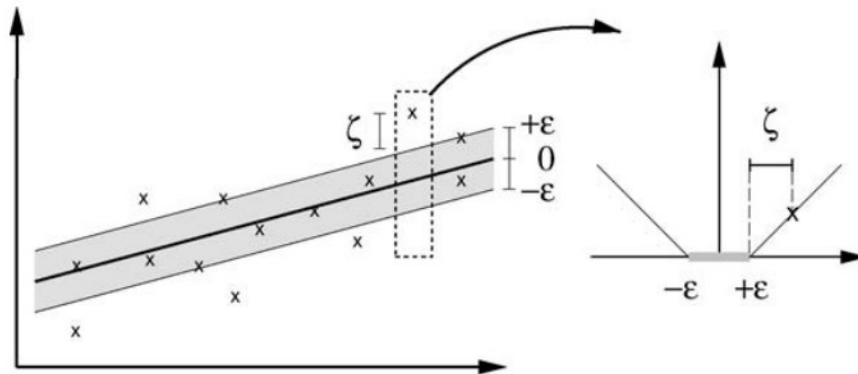


شکل ۹-۲. توابع خطا [۲۱]

تابع خطای درجه دوم در شکل a ۹-۲، با معیار حداقلی خطای مربعات قراردادی مطابقت دارد. تابع خطای شکل b ۹-۲، یک تابع لاپلاسی است که نسبت به تابع خطای درجه دوم در نقاط دورتر حساسیت کمتری دارد. در شکل c ۹-۲، هابر، تابع خطای لجستیک که مشخصات بهینه‌ای در هنگام ناشناخته بودن توزیع اساسی داده‌ها را دارد، پیشنهاد داد. برای پرداختن به این موضوع، وپنیک تابع خطا در شکل d ۹-۲ را به عنوان تقریب تابع خطای هابر که مجموعه‌ای پراکنده از بردارهای پشتیبان را بدست می‌آورد، پیشنهاد می‌کند [۱۶].

^۱Loss Function

مطابق شکل ۱۰-۲، فقط نقاطی که در محدوده $+\varepsilon$ و $-\varepsilon$ قرار دارد در یادگیری نقش داشته و دارای خطایی برابر ζ هستند. نقاط خارج از محدوده مشخص شده، تعداد بردارهای پشتیبان در حل مسئله را تشکیل می‌دهند [۲۲].



شکل ۱۰-۲. تابع خطای ماشین بردار پشتیبان خطی [۲۲]

هدف از پیدا کردن یک رگرسیون خطی

$$f(x) = w^T x + \theta \quad (۱۲-۲)$$

است، به طوری که

$$E_1(w) = \frac{1}{2} \|w\|^2 + C \sum_{p=1}^N \|y_p - f(x_p)\|_{\varepsilon} \quad (۱۳-۲)$$

به حداقل برسد. با $\|0\|_{\varepsilon}$ ، تابع خطای ε -insensitive بصورت زیر تعریف می‌شود.

$$\|x\|_{\varepsilon} = \max\{0, \|x\| - \varepsilon\} \quad (۱۴-۲)$$

که $\varepsilon > 0$ و $C > 0$ از پیش تعیین شده هستند.

این مسئله می‌تواند به یک مسئله دو جمله‌ای تبدیل شود:

$$\text{Minimize } E_2(w, \xi, \zeta) = \frac{1}{2} \|w\|^2 + C \sum_{p=1}^N (\xi_p + \zeta_p) \quad (۱۵-۲)$$

منوط به

$$(w^T x_p + \theta) - y_p \leq \varepsilon + \xi_p \quad (۱۶-۲)$$

$$y_p - (w^T x_p + \theta) \leq \varepsilon + \zeta_p \quad (۱۷-۲)$$

$$\xi_p \geq 0, \quad \zeta_p \geq 0 \quad (۱۸-۲)$$

است. همچنین $p = 1, \dots, N$ می باشد.

زمانی که خطا کمتر از ε باشد، مقدار متغیر کمکی ξ_p و ζ_p صفر می گردد.

با تعمیم رگرسیون خطی به تخمین رگرسیون برپایه کرنل و استفاده از روش ضریب لاگرانژ،

مسئله بهینه سازی زیر بدست می آید.

$$E_{SVR}(\alpha, \beta) = \frac{1}{2} \sum_{p,i=1}^N (\alpha_p - \beta_p)(\alpha_i - \beta_i) \kappa(x_p, x_i) + \sum_{p=1}^N (\alpha_p - \beta_p) y_p + \sum_{p=1}^N (\alpha_p - \beta_p) \varepsilon \quad (19-2)$$

بنابراین

$$\sum_{p=1}^N (\alpha_p - \beta_p) = 0 \quad (20-2)$$

$$0 \leq \alpha_p, \beta_p \leq C \quad (21-2)$$

برای $p = 1, \dots, N$ ، زمانی که ضریب لاگرانژ α_p و β_p به ترتیب با ۲-۱۷ و ۲-۱۸ مطابقت دارد.

خروجی SVM باعث ایجاد رگرسیون

$$u(x) = f(x) = \sum_{p=1}^N (\alpha_p - \beta_p) \kappa(x_p, x) + \theta \quad (22-2)$$

می شود. در این رابطه θ با استفاده از شرایط مرزی قابل حل می باشد.

رگرسیون بردار پشتیبان به دلیل معرفی ε -insensitive، روشی قوی می باشد. متغیر ε بر تعداد

بردارهای پشتیبان تأثیر گذاشته و پیچیدگی مدل را کنترل می کند [۷].

۲-۷-۲ حالت غیرخطی رگرسیون بردار پشتیبان

رگرسیون بردار پشتیبان با استفاده از یک پیش پردازش ساده بر روی داده‌های آموزش، قابلیت حل مسائل غیرخطی را دارد. بدیهی است مسائل غیرخطی پیچیدگی بیشتری نسبت به مسائل خطی دارد. بنابراین یادگیری روابط غیرخطی با استفاده از یک ماشین خطی نیازمند نمایش متفاوت داده‌ها به صورت ویژگی‌های غیرخطی می‌باشد.

از خصوصیات مهم در یادگیری ماشین خطی، قابلیت بیان دوگانه در آن‌ها می‌باشد که منجر به ضرب داخلی بین داده‌ها می‌گردد. با بهره‌گیری از خاصیت بیان دوگانه در حالت غیرخطی، تخمین تابع با استفاده از ضرب داخلی بردار $\phi(x)$ برای مقدار x ، ورودی و بردارهای ورودی آموزش نگاشت داده شده، امکان پذیر خواهد بود.

اکثر مسائلی که توسط رگرسیون بردار پشتیبان مدلسازی می‌گردد غیرخطی است. انتخاب تابع کرنل، متناسب با شرایط داده‌ها از اهمیت ویژه‌ای برخوردار است. در تحقیقات انجام شده نتایج بدست آمده از تابع RBF نشان دهنده کارایی بهتر این تابع نسبت به سایر توابع می‌باشد [۲۳].

برخی از توابع کرنل در جدول ۱-۲ آورده شده است.

جدول ۱-۲. برخی از توابع کرنل پرکاربرد [۲۴]

$\langle x, x' \rangle$		کرنل خطی
$(a\langle x, x' \rangle + b)^p$	a, b, p	کرنل چند جمله ای
$e^{-\sigma \ x - x'\ ^2}$	σ	کرنل RBF
$\tanh(a + \langle x, x' \rangle + b)$	a, b	کرنل تانژانت هایپربولیک

۸-۲ سختی آب

اغلب منابع کوچک آب همانند برخی از منابع آب‌های سطحی بزرگ با استفاده از آب‌های زیرزمینی با سختی قابل توجهی روبرو هستند. غلظت کلسیم بیش از ۱۰۰ میلی‌گرم در لیتر در منابع طبیعی آب، به ویژه آب‌های زیرزمینی معمول است. منیزیم در آب‌های زیرزمینی معمولاً دارای غلظت کمتری می‌باشد (از مقدار ناچیز تا حدود ۵۰ میلی‌گرم در لیتر و بندرت بالاتر از ۱۰۰ میلی‌گرم در لیتر)، بنابراین معمولاً سختی بر پایه کلسیم غالب می‌باشد. میزان جذب منیزیم از آب، براساس مصرف ۲ لیتر آب در هر روز برای یک انسان بالغ به ترتیب حدود ۲/۳ میلی‌گرم و ۵۲/۱ میلی‌گرم در آب نرم^۱ و آب سخت^۲ گزارش شده است. آب‌های طبیعی و تصفیه شده دارای طیف وسیعی از مواد معدنی، از سطح بسیار کم در آب باران و آب نرم و رسوب‌زدایی تا سطح متوسط و سطح بسیار زیاد در آب‌های سخت و طبیعی با محتوای کل مواد جامد محلول، هستند.

انتخاب مناسب‌ترین تکنولوژی تصفیه به شرایط محلی بستگی دارد (به عنوان مثال مسائل مربوط به کیفیت آب، جنس لوله کشی، خوردگی). برخی از شیرین سازها در تصفیه خانه مرکزی و برخی دیگر در خانه‌های شخصی انجام می‌شود. اصلاح غلظت کلسیم و منیزیم در آب آشامیدنی به دلایل بهداشتی باید مطابق با الزامات فنی برای توزیع آب بوده و گندزدایی را به خطر نیندازد. براساس شرایط محلی و کمبودهای رایج ممکن است متناسب با مواد معدنی مغذی، تأمین کنندگان آب و مقامات بهداشت عمومی نیاز به اصلاح بیشتر در ترکیب آب آشامیدنی را داشته باشند. مقدار کلسیم و منیزیم کلسیم و منیزیم در آب آشامیدنی، ممکن است برای افرادی که محدودیت در دریافت آن را دارند، بسیار مهم باشد [۲۵].

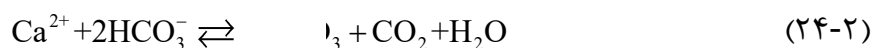
^۱Soft Water

^۲Hard Water

برای بیشتر اهداف عملی، سختی آب را می‌توان به عنوان مجموع غلظت کلسیم و منیزیم نشان داد که با میلی اکی والان در لیتر داده می‌شود:

$$\text{Hardness, eq/L} = 2[\text{Ca}^{2+}] + 2[\text{Mg}^{2+}] \quad (23-2)$$

در رابطه ۲-۲۴، غلظت Ca و Mg برحسب mol/L و ضریب ۲ ماهیت دو ظرفیتی هر دو یون را بیان می‌کند. بطور کلی دو نوع سختی داریم: سختی کربناته شامل HCO_3^- و CO_3^{2-} ، و سختی غیرکربناته که شامل یون‌های دیگر به ویژه Cl^- و SO_4^{2-} می‌باشد. موازنه بین سختی کربناته و غیرکربناته در نرم شدن آب (حذف سختی) و تشکیل رسوب می‌باشد. از آنجا که HCO_3^- در دمای بالا جدا می‌شود، نتیجه گرم شدن آب سخت، تشکیل رسوب ناشی ته‌نشینی HCO_3^- به صورت:



رسوب دریچه لوله‌ها ضرایب انتقال حرارت را کاهش و مقاومت اصطکاکی جریان در لوله‌ها را تغییر می‌دهد [۲۶].

فصل ۳ : مبانی نظری و پیشینه تحقیق

۳-۱ مقدمه

آب سخت بدلیل خطرات زیاد برای سلامتی، علت بسیاری از بیماری‌ها در انسان بوده و صنایع را تحت تاثیر قرار می‌دهد. سختی موجود در آب توسط یون‌های فلزی چند ظرفیتی سنگ‌های رسوبی، نفوذ و رواناب در خاک ایجاد می‌شود. کلسیم و منیزیم دو یون اصلی که در بسیاری از سنگ‌های رسوبی وجود دارد، یون‌های اصلی ایجاد کننده سختی آب هستند. مصرف بیش از حد کلسیم و منیزیم خطرات پوکی استخوان، سنگ کلیه^۱ سرطان روده بزرگ، فشار خون و سکته، بیماری سرخرگ کرونری^۲ و چاقی را می‌تواند افزایش دهد. برخی مطالعات همه‌گیرشناسی و تجربی، کمبود

^۱Nephrolithiasis

^۲Coronary

منیزیم در بیماران دارای فشار خون بالا نشان از ارتباط بین فشار خون و سطح منیزیم می‌باشد. با وجود کاربرد شبکه عصبی در مدلسازی منابع آب و هیدرولوژی سعی شده از آن در بررسی کیفیت منابع آب مورد مطالعه قرار گیرد [۴].

ظرفیت آب برای واکنش با صابون روشی مرسوم برای اندازه‌گیری سختی آب است. آب سخت به صابون بیشتری برای ایجاد کف صابون نیاز دارد. همچنین اکثر مواقع رسوبات قابل توجهی در ظرف ته نشین می‌شود. این ماده توسط انواع یون‌های فلزی چند ظرفیتی محلول (به عنوان مثال آلومینیوم، باریم، آهن، منگنز، استرانسیم و روی) که عمدتاً کاتیون‌های کلسیم و منیزیم هستند، با مشارکت کاتیون‌های دیگر ایجاد می‌گردد. به طور کلی آب حاوی کربنات کلسیم در غلظت‌های زیر ۶۰ میلی‌گرم در لیتر نرم، ۶۰-۱۲۰ میلی‌گرم در لیتر نیمه سخت، ۱۲۰-۱۸۰ میلی‌گرم در لیتر سخت و بیشتر از ۱۸۰ میلی‌گرم در لیتر بسیار سخت می‌باشد.

بیشتر منابع طبیعی سختی در آب، یون‌های فلزی چند ظرفیتی محلول از سنگ‌های رسوبی، تراوش و رواناب از خاک است.

کلسیم و منیزیم علاوه بر ترکیبات معدنی ضروری مواد غذایی رایج، دو یون اصلی هستند که در بسیاری از سنگ‌های رسوبی وجود داشته و متداول‌ترین آن سنگ آهک و گچ می‌باشد. همانطور که در بالا ذکر شد، سهم جزئی در سختی کل آب توسط یون‌های چند ظرفیتی دیگر، مانند آلومینیوم، باریم، آهن، منگنز، استرانسیم و روی نیز انجام می‌شود.

۳-۲ مروری بر چند مطالعه در زمینه مدلسازی کیفی به کمک شبکه عصبی

پیش بینی تراز آب‌های زیرزمینی برای استفاده و مدیریت منابع آب زیرزمینی مهم می باشد. در این تحقیق، شيامین و همکاران در سال ۲۰۰۲ از شبکه عصبی مصنوعی (ANN) برای پیش بینی تراز آب‌های زیرزمینی در آبخوان Dawu در شرق چین استفاده کردند. تغییر ماهانه تراز آب‌های زیرزمینی اولین قدم برای آنالیز همبستگی اتوماتیک است نشان دهنده وابستگی آن به زمان می باشد. مدل ARANN^۱ و RARANN^۲ با استفاده از الگوریتم پس انتشار برای پیش‌بینی تراز آب‌های زیرزمینی استفاده شد. داده‌های مورد استفاده در این مدلسازی از سال ۱۹۸۸ تا ۱۹۹۸ برای آموزش و آزمون شبکه استفاده شد. نتایج بدست آمده نشان دهنده کارایی بهتر مدل RARANN نسبت به مدل ARANN می باشد. نتایج نشان دهنده توانایی این مدل در پیش بینی نوسانات تراز آب‌های زیرزمینی در منطقه مورد نظر برای شرایط مشابه در آینده می باشد [۲۷].

گولوکا^۳ و همکاران در این مطالعه از یک شبکه عصبی پیشرو^۴ برای پیش‌بینی غلظت آلودگی آفت‌کش چاه‌های تحت نظارت استفاده کردند. داده‌ها شامل شاخص‌های طبقه‌ای مواد عمق آبخوان، طبقه جداسازی سموم دفع آفات، حساسیت آبخوان به آلودگی سموم دفع آفات، زمان (ماه) نمونه برداری، عمق چاه، عمق آب تا سطح زمین و فاصله حرکت در منطقه اشباع می باشد (یعنی فاصله از سطح زمین تا میانه صفحه چاه). خروجی شبکه عصبی مجموع آلودگی سموم آفت‌کش در چاه است. نتایج نشان می‌دهد چاه‌های کم عمق در معرض آلودگی بیشتری نسبت به چاه‌های عمیق قرار دارند. همچنین میزان سموم در ماه‌های پر کار ۲/۵ تا ۳ برابر بیشتر از سایر ماه‌ها می باشد. پارامترها شامل

^۱Auto-Regression

^۲Regression-Auto-Regression

^۳Goloka B. Sahoo

^۴Feed-Forward back-propagation Neural Network (BPNN)

دو پارامتر اصلی (مواد در عمق آبخوان و نوع شستشوی آفت کش ها) و پنج پارامتر جانبی (مانند عمق و زمان جمع آوری نمونه) می باشد. پارامترهای اصلی به تنهایی قادر به تحلیل نبوده و برای راندمان بالاتر نیاز به پارامترهای جانبی بوده است. این عملکرد در اعداد بدست آمده از R ، E ، ME ، $RMSE$ و بروز نوع آفت کش قابل مشاهده می باشد. شبکه عصبی پیشرو نشان داد که داده های جانبی قدرت پیش بینی بیشتری نسبت به داده های اصلی دارند. نتایج به محققان کمک می کند تا به بهبود پارامترهای نقشه حساسیت سفره آب به آلودگی سموم دفع آفات شناسایی کنند [۲۸].

گمیتزی^۱ و همکاران در مطالعه خود در سال ۲۰۰۸، مهمترین پارامترهای کنترل آلودگی آب های زیرزمینی توسط نیترات را تعیین کردند. پارامترها شامل هدایت هیدرولیکی آبخوان، عمق آبخوان، کاربری اراضی، نفوذپذیری خاک و نسبت درشت دانه در منطقه غیر اشباع می باشند که در محیط GIS با یک مقیاس مشترک استاندارد سازی می شوند. به منظور دسته بندی پتانسیل آلودگی نیترات آب های زیرزمینی منطقه مورد بررسی از پرسپترون چندلایه^۲ طبقه بندی شبکه عصبی با الگوریتم پس انتشار استفاده شد. از تحلیل آزمایش ۲۱۴ محل آلودگی نیترات آب های زیرزمینی، دقت پیش بینی برای نمونه های آزمایش ۸۶٪، مجموع کل نمونه ها ۷۴٪ و نمونه آزمایشگاهی ۷۱٪ بدست آمد [۲۹].

خامیس^۳ و همکاران در سال ۲۰۱۲ یک مدل جدید برای تخمین غلظت آلودگی نیترات آب های زیرزمینی در آبخوان نوار غزه^۴ را ارائه دادند. مجموعه ای از شش متغیر توصیفی برای ۱۳۹ نمونه چاه که دارای اهمیت برای مدل شبکه عصبی بودند، استفاده شد. بهترین روش یافت شده برای مدلسازی نیترات، روش چند لایه بوده که دارای شش گره ورودی و چهار گره پنهان می باشد. متغیرهای ورودی

^۱A. Gemitzi

^۲Multi-Layer Perceptrons (MLP)

^۳khamis

^۴Gaza strip

عبارتند از بار نیتروژن، میزان تراکم ساکنین در شعاع ۵۰۰ متری چاه، عمق چاه، طول، دبی و سرعت نفوذ می‌باشد. بر اساس مدل شبکه عصبی، آلودگی نیترات آب‌های زیرزمینی به هیچ یک از عوامل به صورت انفرادی بستگی نداشته و به ترکیب آن‌ها وابسته است [۳۰].

آلودگی آرسنیک آب‌های زیرزمینی تاثیر منفی بر روی سلامت انسان دارد. برای حل این مشکل، پیش‌بینی میزان آلودگی برای جلوگیری از صدمه احتمالی به انسان توسط مکتبول^۱ و همکاران در سال ۲۰۱۳ انجام شد. در این تحقیق نقشه خطر آلودگی آرسنیک با استفاده از درخت تصمیم و GIS تهیه شده است. عواملی که ممکن است با آلودگی آرسنیک ارتباط داشته باشد عبارتند از زمین‌شناسی، واکنش خاک (pH)، تراز، ماده آلی خاک، آهن خاک و میزان نزدیکی به رودخانه به عنوان متغیرهای پیش‌بینی کننده انتخاب شدند. مدل پیش‌بینی با استفاده از الگوریتم طبقه‌بندی داده کاوی در MATLAB با استفاده از ۸۰٪ داده‌های آموزشی تهیه شده است. بهترین درخت در اصلاح سطح ۶ با روش اعتبارسنجی متقابل ۸۷/۲۵ یافت شد. بهترین درخت به دست آمده از فرآیند روش اعتبارسنجی متقابل برای ایجاد نقشه خطر آرسنیک در محیط GIS استفاده شد. نقشه ریسک توسعه داده شده به مقدار ۸۷/۹٪ نزدیک به واقعیت می‌باشد. این تحقیق نشان می‌دهد با حمایت از فرآیند ژئوشیمیایی تراز سطح یک منطقه، میزان تجمع آرسنیک در آب‌های زیرزمینی قابل تعیین می‌باشد [۳۱].

جنگل تصادفی^۲ یک روش یادگیری ماشین است که در مقایسه با تکنیک‌های شناسایی الگوی متعارف به ندرت در مطالعات آب استفاده شده است. مزایای اصلی روش RF عبارتند از: ماهیت غیر پارامتری آن، دقت پیش‌بینی بالا و قابلیت تعیین اهمیت متغیر می‌باشد. آخرین ویژگی بیان شده می‌تواند برای درک بهتر نقش فردی و اثر ترکیبی متغیرهای توضیحی در محافظت و قرار گرفتن یک آلاینده در معرض آب‌های زیرزمینی مورد استفاده قرار گیرد. در این تحقیق عملکرد رگرسیون RF

^۱Md. Moqbul Hossain

^۲Random Forest(RF)

برای مدلسازی پیش بینی آلودگی نیترات بر اساس ارزیابی آسیب پذیری ذاتی و خاص آبخوان Vega de Granada مورد بررسی قرار گرفت. یک پایگاه داده جامع GIS متشکل از ۲۴ پارامتر مرتبط به مشخصات هیدرولوژیکی، نیروی محرک، متغیرهای سنجش از دور و متغیرهای فیزیکی و شیمیایی اندازه گیری شده در محل، بعنوان ورودی برای ساخت مدل های مختلف پیش بینی آلودگی نیترات استفاده شده است. اهمیت اندازه گیری به روش RF برای تعریف پیش بینی آلودگی نیترات در آب های زیرزمینی که باعث ایجاد منابع آلودگی می شود قابل استفاده می باشد. پتانسیل RF برای ایجاد نقشه آسیب پذیری در برابر آلودگی نیترات با توجه به معیارهای مربوط به تغییرات پارامترهای الگوریتم و دقت نقشه ها ارزیابی شد. عملکرد RF در مقایسه با روش رگرسیون لجستیک با استفاده از میزان راندمان مختلف برای اطمینان از قابلیت تعمیم آن ها ارزیابی گردید. نتایج پیش بینی توانایی RF در ساخت مدل های دقیق با قابلیت های پیش بینی قوی را نشان می دهد [۳۲].

راحله عربگل و همکاران مدل برای پیش بینی غلظت نیترات در سفره دشت اراک با استفاده از مجموعه ای از متغیرهای کیفیت آب زیرزمینی که به راحتی قابل اندازه گیری هستند، ارائه کردند. دما، هدایت الکتریکی، عمق آب زیرزمینی، میزان جامدات محلول، میزان اکسیژن محلول، اسیدیته، کاربری اراضی و موقعیت مکانی چاه، به عنوان متغیرهای ورودی بکار گرفته شده است. برای این منظور از داده های کیفی جمع آوری شده از ۴۰ چاه در ۴ فصل متوالی به مدت یک سال استفاده شد. مدل ماشین بردار پشتیبان غلظت نیترات را در مجموعه داده های مرحله آزمایش و آزمون با همبستگی معقول (به ترتیب ۰/۹۲ و ۰/۸۷) با مقادیر اندازه گیری و خطای میانگین کمترین مجذور به ترتیب ۰/۰۸۶ و ۰/۱۱۱۱ پیش بینی کرد. به طور کلی، نتایج نشان داد مدل ماشین بردار پشتیبان میتواند به عنوان روشی سریع، مطمئن و مقرون به صرفه برای ارزیابی و پیش بینی کیفیت آب زیرزمینی مورد استفاده قرار گیرد [۳۳].

در مطالعات انجام شده توسط رحیم برزگر و همکاران به مطالعه آب‌های زیرزمینی در منطقه ماکو شمال غربی ایران پرداختند. هدف از این مطالعه ارزیابی توانایی ماشین یادگیری افراطی^۱ برای پیش‌بینی میزان آلودگی غلظت فلوراید در آب‌های زیرزمینی در مقایسه با مدل پرسپترون چند لایه و ماشین بردار پشتیبان می‌باشد. برای این منظور از ۱۴۳ نمونه آب در یک دوره ۵ ساله از ۲۰۰۴ تا ۲۰۰۸ جمع آوری شد. نمونه‌ها از نظر هدایت الکتریکی، PH، یون‌های شیمیایی اصلی و فلوراید اندازه گیری و تحلیل شدند. برای توسعه مدل، مجموعه داده‌های شامل Na^+ , K^+ , Ca^{2+} و HCO_3^- به عنوان غلظت ورودی و غلظت فلوراید به عنوان غلظت خروجی به دو زیر مجموعه تقسیم شد؛ آموزش (۸۰٪ داده) و آزمون (۲۰٪ داده) به روش اعتبارسنجی متقابل انجام شد. نتایج بدست آمده از مدل ماشین یادگیری افراطی نشان می‌دهد R^2 برابر ۰/۹۲۱، NSC برابر ۰/۹۰۷۱، RMSE برابر ۰/۵۶۳۸ میلی‌گرم در لیتر و MABE برابر ۰/۴۶۳۵ میلی‌گرم در لیتر برای داده‌های آموزش می‌باشد. نتایج نشان می‌دهد، مدل یادگیری افراطی عملکرد بهتری نسبت به پرسپترون چند لایه و مدل ماشین بردار پشتیبان برای پیش‌بینی آلودگی فلوراید را دارد [۱۱].

استفاده بیش از حد از کودهای شیمیایی به ویژه کود نیتروژن برای افزایش محصول و تصفیه نامناسب و انتقال فاضلاب شهری و صنعتی به عنوان عوامل افزایش میزان نیترات در آب‌های زیرزمینی در این منطقه می‌باشد. بنابراین، بررسی آلودگی نیترات به عنوان یکی از مهمترین مشکلات زیست محیطی در آب‌های زیرزمینی ضروری است. مطالعه انجام شده توسط مدلسازی و برآورد آلودگی نیترات در کیفیت آب‌های زیرزمینی منطقه حاشیه رودخانه زاینده رود اصفهان با استفاده از شبکه‌های عصبی مصنوعی مورد بررسی قرار گرفت. ۱۰۰ حلقه چاه (۷۷ مورد چاه کشاورزی، ۱۳ چاه آب آشامیدنی و ۱۰ چاه باغ) در منطقه حاشیه رودخانه زاینده رود انتخاب شد که پس از آزمایش‌های مکرر، شبکه‌ای با یک لایه پنهان و ۱۹ نورون، کمترین خطا را در روند آموزش، آزمون و اعتبارسنجی

^۱Extreme Learning Machine

شبکه ایجاد گردید. مدل‌های شبکه عصبی برای بررسی پارامترهای کیفیت آب نیز قابل استفاده هستند[۳۴].

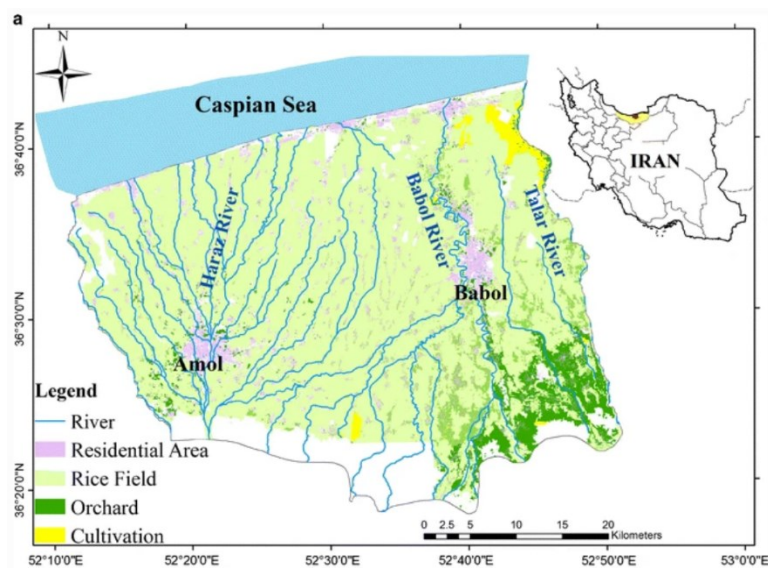
هانان درویش^۱ و همکاران در سال ۲۰۱۷ روشی (GIS /ANN) مناسب برای مسائل شبیه سازی در مقیاس بزرگ و بلند مدت پیشنهاد کردند که در پیش‌بینی مکانی نیترات‌ها در آبخوان چاک در بتون (شمال فرانسه) استفاده شد. توزیع مکانی بدست آمده غلظت نیترات با شبکه عصبی و سیستم اطلاعات مکانی و نتایج مدل‌سازی عددی (MT3D) نشان می‌دهد، مدل‌سازی شبکه عصبی واقعی‌تر از مدل‌سازی MT3D در سال ۲۰۰۳ و ۲۰۰۴ می‌باشد. نتایج پتانسیل مدل‌سازی شبکه عصبی با استفاده از داده‌های توزیع مکانی هیدروژنولوزیکی برای مدیریت آب‌های زیرزمینی آلودگی نیترات را نشان می‌دهد. از نظر مدیریت آب‌های زیرزمینی، مدل‌سازی در GIS و شبکه عصبی، روشی جایگزین برای به دست آوردن نتایج سریع و رضایت بخش می‌باشد[۳۵].

مطالعات انجام شده توسط عیسی زاده و همکاران نشان دهنده کارایی بهتر SVM نسبت به شبکه عصبی می‌باشد. در این مطالعه پارامترهای کیفی آب‌های زیرزمینی استان گیلان، واقع در شمال ایران توسط شبکه عصبی و ماشین بردار پشتیبان مدل‌سازی شد. اطلاعات مورد نیاز از ۱۴۰ چاه برای سال‌های ۲۰۰۲ تا ۲۰۱۴ جمع‌آوری شد. پنج متغیر شامل مختصات X و Y چاه‌های مشاهداتی، فاصله هر چاه تا خط ساحلی، میانگین عمق بارندگی ۶ ماهه و تراز آب‌های زیرزمینی در روزهای نمونه برداری کیفیت آب به عنوان متغیرهای ورودی در نظر گرفته شد. علاوه بر این، نه متغیر بعنوان متغیرهای ورودی کمکی در نظر گرفته شد. متغیرهای کمکی آب‌های زیرزمینی منطقه شامل هدایت الکتریکی، غلظت سدیم و غلظت سولفات است که با استفاده از ANN و SVM با ترکیب متغیرهای ورودی مختلف مورد ارزیابی قرار گرفت. نتایج نشان می‌دهند کارایی SVM و ANN زمانی مناسب است که تنها متغیر ورودی اصلی، محل چاه باشد[۱۲].

^۱Hanan Darwishe

۳-۳ منطقه مورد مطالعه

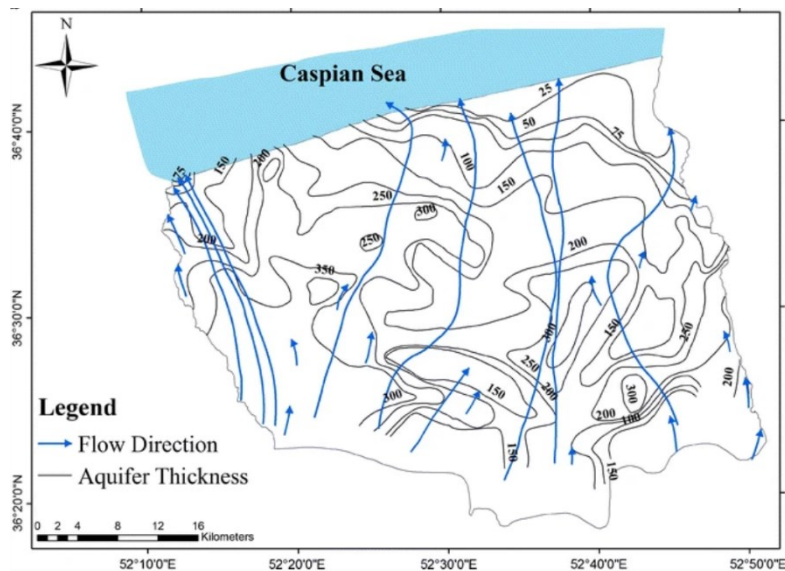
منطقه مورد مطالعه، دشت آمل-بابل در شمال ایران واقع در استان مازندران می‌باشد. مساحت تقریبی این منطقه 1822 Km^2 است که بین $36^{\circ}16'$ و $36^{\circ}47'$ عرض شمالی و $52^{\circ}9'$ و $52^{\circ}59'$ طول شرقی قرار دارد [۳۶]. کشاورزی اصلی ترین فعالیت اقتصادی است که حدود ۸۰ درصد مناطق دشت را پوشش می‌دهد. محصول اصلی این منطقه برنج بوده که تقریباً ۶۲ درصد (1200 Km^2) از زمین برای کشت برنج استفاده می‌شود. همچنین، ۱۲ درصد از زمین شامل باغ مرکبات و ۱۰ درصد را محصولاتی مانند گندم، جو، لوبیا و ذرت پوشش می‌دهند (شکل ۳-۱). تمام آب خانگی و حدود ۳۰ درصد از آبیاری کشاورزی از آب‌های زیرزمینی تامین می‌شود. منطقه مورد نظر از نظر ژئومورفولوژی یک دشت آبرفتی مسطح است (عمدتاً زاویه شیب کمتر از ۱ درصد) که توسط دریای خزر در شمال و ارتفاعات البرز در جنوب احاطه شده است [۳۷].



شکل ۳-۱. نقشه کاربری اراضی بابل-آمل [۳۷]

آبخوان آبرفتی متشکل از یک آبخوان محدود است که تا حدود ۹۴٪ دشت را تشکیل می‌دهد. ضخامت آبخوان از غرب و جنوب غربی در حال افزایش و از سمت شمال در حال کاهش است که محدوده تغییرات آن از ۱۰ تا ۲۰۰ متر می‌باشد (شکل ۳-۲). تخلیه جریان منطقه‌ای آب زیرزمینی از

قسمت جنوب (کوه البرز) به شمال (دریای خزر) می‌باشد. بطور طبیعی این دبی از نفوذ بارش و برف در ارتفاعات البرز حاصل شده و تخلیه آن به سمت رودخانه‌ها، پهنه آبی و نهایتاً به دریای خزر رخ می‌دهد. میزان بالای دبی رودخانه‌های بابل و هراز نفوذ آب دریا را به قسمت منطقه ساحلی در ناحیه شمال شرقی دشت محدود می‌کند. این میزان از دبی آب مانده شور را از طریق لایه‌های رسوبی شسته و از نفوذ آب دریا به آبخوان جلوگیری می‌کند. از طرفی، در شمال شرقی میزان دبی پایین رودخانه تالار قدرت کافی شستشوی آب مانده شور را ندارد. بنابراین آب شور بین لایه‌های رسوب باقی مانده و کیفیت آب زیرزمینی را در این منطقه کاهش می‌دهد [۳۷].



شکل ۳-۲. نقشه ضخامت آبخوان [۳۷]

فصل ۴: نتایج و بحث

۴-۱ مقدمه

استفاده از شبکه عصبی به دلیل قابلیت تعمیم پذیری در مقابل سایر روش‌ها، روشی نو در مدلسازی غلظت مواد در آب‌های زیرزمینی می‌باشد. از جمله قابلیت‌های شبکه عصبی می‌توان به پیش‌بینی، پردازش تصویر، تشخیص چهره و تشخیص صدا اشاره نمود. به منظور پیش‌بینی الگوی پراکندگی سختی موقت و دائم آب‌های زیرزمینی دشت آمل-بابل از ۸۵۵ چاه منتخب استفاده شد که در فصل پاییز و بهار در فاصله زمانی پاییز ۱۳۹۲ تا پاییز ۱۳۹۸ نمونه برداری شد. به منظور صحت سنجی داده‌های بدست آمده در شبکه عصبی پیشرو از ماشین بردار پشتیبان جهت مقایسه استفاده گردید. پس از مدلسازی و بدست آوردن پارامترهای بهینه، جهت نمایش پراکندگی داده‌های پیش‌بینی شده از نرم‌افزار GIS استفاده شد.

۲-۴ مشخصات پارامترهای ورودی و خروجی

همانطور که در این تحقیق اشاره شد، از داده‌های موجود و استخراج شده از ۸۵۵ نمونه (در دو فصل) برای ساخت مدلی بهینه به منظور تخمین غلظت سختی موقت و دائم استفاده شده است. در تقسیم بندی داده‌ها به منظور آماده سازی مدل شبکه عصبی بصورت تصادفی، از کد MATLAB برای این کار استفاده گردید. بطوری که ۸۰ درصد از داده‌ها جهت آموزش و ۲۰ درصد باقی مانده برای آزمون و اعتبارسنجی استفاده شده است. به عبارت دیگر می‌توان گفت از داده‌های استخراج شده ۶۸۴ نمونه بصورت تصادفی از کل داده‌ها برای آموزش مدل و سایر آن برای آزمون و اعتبارسنجی بکارگرفته شده است. پارامترهای استفاده شده در این تحقیق عبارتند از SO_4^{2-} ، Cl^- ، TH ، PH ، EC ، SAR و K^+ و NO_3^- که به عنوان داده‌های ورودی و سختی موقت و دائم در چاه‌های مورد نظر به عنوان خروجی مدل در نظر گرفته شد.

جدول ۱-۴. اطلاعات آماری کیفیت آب های زیرزمینی

پارامترهای ورودی	واحد	حداکثر	حداقل	میانگین
SO_4^{2-}	Mg/L	۱۲/۱۰	۰/۱۰	۶/۱۰
Cl^-	Mg/L	۳۱/۹۰	۰/۲۰	۱۶/۰۵
NO_3^-	Mg/L	۱۳۹/۵۰	۰/۰۰	۶۹/۷۵
PH	--	۱۰/۶۰	۶/۳۰	۸/۴۵
TH	Meq/L	۱۹/۲۰	۱/۳۰	۱۰/۲۵
SAR	--	۹/۷۱	۰/۰۲	۴/۸۷
EC	$\mu S/cm$	۴۱۸۰	۲۰۰	۲۱۹۰
K^+	Meq/L	۰/۱۴	۰/۰۱	۰/۰۸
پارامتر خروجی				
سختی موقت	Mg/L	۹۵۰	۱۰	۳۱۳/۶۳
سختی دائم	Mg/L	۴۶۰	۰	۵۱/۷۳

مطابق جدول ۴-۲، آنالیز همبستگی انجام شده برای پارامترهای انتخابی نشان دهنده عدم ارتباط

معنادار بین پارامترهای مورد نظر می‌باشد. ضریب همبستگی عددی بین ۱ تا -۱ است. ضریب بین ۰ تا ۱ به معنی داشتن همبستگی مثبت و هرچه این ضریب به ۱ نزدیک‌تر باشد همبستگی قوی‌تر می‌باشد. همبستگی مثبت یعنی با افزایش نمره یک متغیر، نمره متغیر دیگر نیز افزایش می‌یابد. ضریب همبستگی بین ۰ تا -۱ به معنی داشتن همبستگی منفی بین دو متغیر است و هرچه عدد به -۱ نزدیک‌تر باشد یعنی همبستگی منفی قوی‌تر است. همبستگی منفی یعنی با کاهش نمره یک متغیر، نمره متغیر دیگر کاهش می‌یابد.

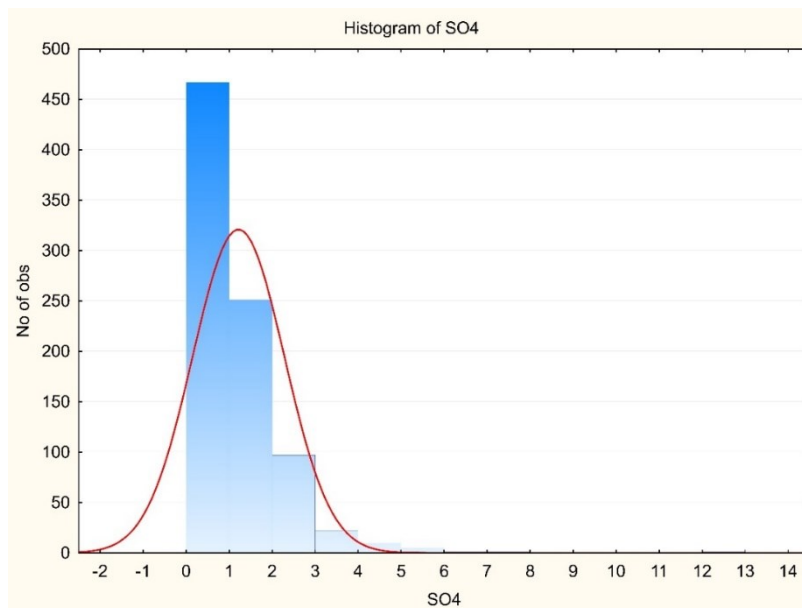
		so4	Cl	NO3	PH	TH	SAR	EC	K	سختی دائم
so4	Pearson Correlation	1	.194**	.149**	-.193**	.433**	.233**	.452**	.399**	.600**
	Sig. (2-tailed)	0	0	0	0	0	0	0	0	0
	N	855	855	855	855	855	855	855	855	855
Cl	Pearson Correlation	.194**	1	-0.055	-0.006	.344**	.831**	.768**	.450**	.560**
	Sig. (2-tailed)	0		0.11	0.869	0	0	0	0	0
	N	855	855	855	855	855	855	855	855	855
NO3	Pearson Correlation	.149**	-0.055	1	-.123**	.245**	-.084*	.114**	.147**	.118**
	Sig. (2-tailed)	0	0.11		0	0	0.014	0.001	0	0.001
	N	855	855	855	855	855	855	855	855	855
PH	Pearson Correlation	-.193**	-0.006	-.123**	1	-.452**	.129**	-.268**	-.366**	-.164**
	Sig. (2-tailed)	0	0.869	0		0	0	0	0	0
	N	855	855	855	855	855	855	855	855	855
TH	Pearson Correlation	.433**	.344**	.245**	-.452**	1	0.059	.765**	.758**	.561**
	Sig. (2-tailed)	0	0	0	0		0.085	0	0	0
	N	855	855	855	855	855	855	855	855	855
SAR	Pearson Correlation	.233**	.831**	-.084*	.129**	0.059	1	.597**	.367**	.242**
	Sig. (2-tailed)	0	0	0.014	0	0.085		0	0	0
	N	855	855	855	855	855	855	855	855	855
EC	Pearson Correlation	.452**	.768**	.114**	-.268**	.765**	.597**	1	.756**	.562**
	Sig. (2-tailed)	0	0	0.001	0	0	0		0	0
	N	855	855	855	855	855	855	855	855	855
K	Pearson Correlation	.399**	.450**	.147**	-.366**	.758**	.367**	.756**	1	.382**
	Sig. (2-tailed)	0	0	0	0	0	0	0		0
	N	855	855	855	855	855	855	855	855	855

		so4	Cl	NO3	PH	TH	SAR	EC	K	سختی موقت
so4	Pearson Correlation	1	.194**	.149**	-.193**	.433**	.233**	.452**	.399**	.112**
	Sig. (2-tailed)		0	0	0	0	0	0	0	0.001
	N	855	855	855	855	855	855	855	855	855
Cl	Pearson Correlation	.194**	1	-0.055	-0.006	.344**	.831**	.768**	.450**	0.033
	Sig. (2-tailed)	0		0.11	0.869	0	0	0	0	0.338
	N	855	855	855	855	855	855	855	855	855
NO3	Pearson Correlation	.149**	-0.055	1	-.123**	.245**	-.084*	.114**	.147**	.215**
	Sig. (2-tailed)	0	0.11		0	0	0.014	0.001	0	0
	N	855	855	855	855	855	855	855	855	855
PH	Pearson Correlation	-.193**	-0.006	-.123**	1	-.452**	.129**	-.268**	-.366**	-.433**
	Sig. (2-tailed)	0	0.869	0		0	0	0	0	0
	N	855	855	855	855	855	855	855	855	855
TH	Pearson Correlation	.433**	.344**	.245**	-.452**	1	0.059	.765**	.758**	.824**
	Sig. (2-tailed)	0	0	0	0		0.085	0	0	0
	N	855	855	855	855	855	855	855	855	855
SAR	Pearson Correlation	.233**	.831**	-.084*	.129**	0.059	1	.597**	.367**	-.094**
	Sig. (2-tailed)	0	0	0.014	0	0.085		0	0	0.006
	N	855	855	855	855	855	855	855	855	855
EC	Pearson Correlation	.452**	.768**	.114**	-.268**	.765**	.597**	1	.756**	.541**
	Sig. (2-tailed)	0	0	0.001	0	0	0		0	0
	N	855	855	855	855	855	855	855	855	855
K	Pearson Correlation	.399**	.450**	.147**	-.366**	.758**	.367**	.756**	1	.654**
	Sig. (2-tailed)	0	0	0	0	0	0	0		0
	N	855	855	855	855	855	855	855	855	855

جدول ۴-۲. ضریب همبستگی خطی (R) بین سختی دائم و سختی موقت و پارامترهای ورودی

۱-۲-۴ سولفات (Sulfate)

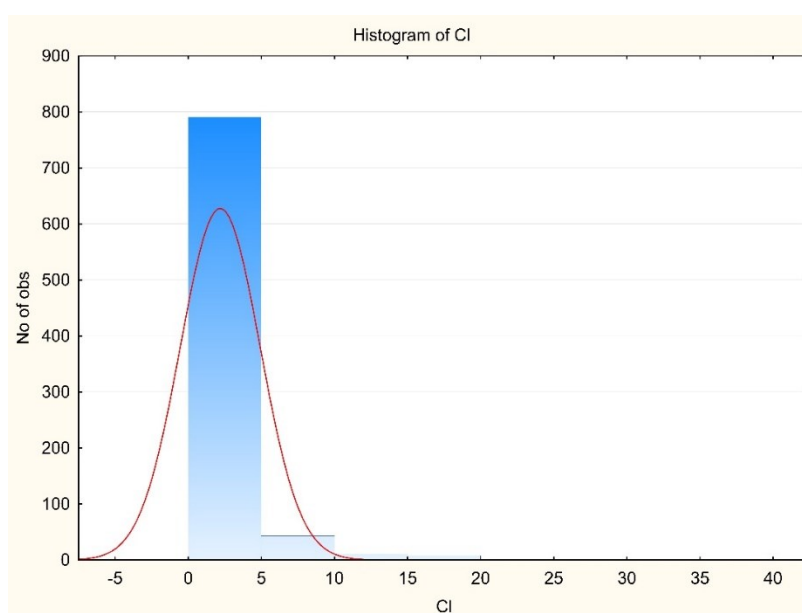
سولفات ها به طور طبیعی در بسیاری از مواد معدنی وجود داشته و به طور عمده در صنایع شیمیایی مورد استفاده تجاری قرار می گیرند. بالاترین سطح از سولفات معمولاً در آب های زیرزمینی از منابع طبیعی می باشد. به طور کلی، متوسط مصرف روزانه سولفات از آب آشامیدنی، هوا و غذا تقریباً ۵۰۰ میلی گرم است که اصلی ترین منبع آن غذا است. با این حال، در مناطقی که آب آشامیدنی حاوی مقادیر زیادی سولفات هستند، ممکن است منبع اصلی مصرف باشد. به دلیل عدم تشخیص سولفات در آب آشامیدنی در داده های سطحی موجود، اثرات نامناسبی بر روی سلامت انسان دارد. در شکل ۳-۴ هیستوگرام داده های موجود برای سولفات قابل مشاهده می باشد [۳۸].



شکل ۳-۴. هیستوگرام فراوانی داده های سولفات

۲-۲-۴ کلرید (Chloride)

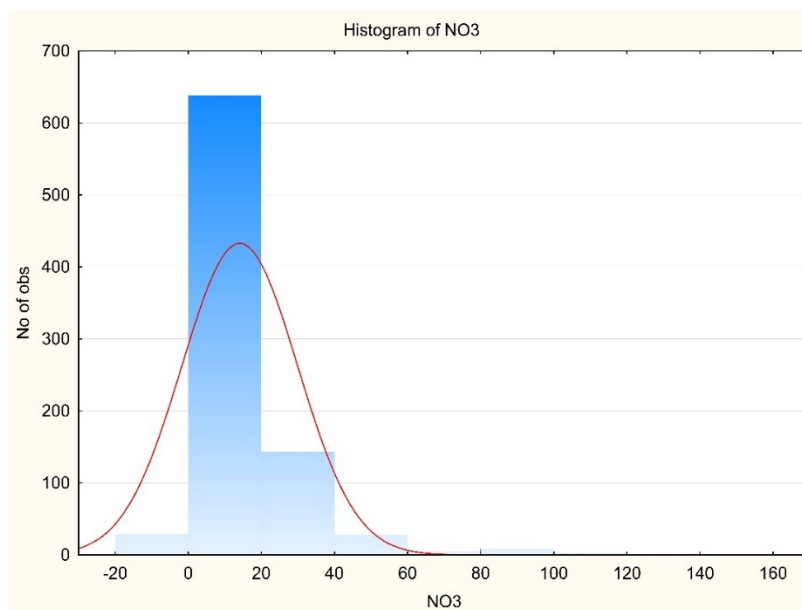
کلرید موجود در آب آشامیدنی از منابع طبیعی، فاضلاب و پساب‌های صنعتی، رواناب شهری حاوی نمک ضد یخ و پیشروی آب دریا سرچشمه می‌گیرد. منبع اصلی قرار گرفتن انسان در معرض کلرید، افزودن نمک به غذا است که میزان دریافت از این منبع معمولاً بیشتر از میزان آب آشامیدنی می‌باشد. غلظت بیش از حد کلرید بسته به قلیایی بودن آب، باعث افزایش خوردگی فلزات در سیستم توزیع می‌شود. دستورالعمل مبتنی بر مقدار مجاز برای کلرید در آب آشامیدنی ارائه نشده است با این حال غلظت بیشتر از 250 mg/L طعم آب تحت تاثیر قرار می‌گیرد [۳۸].



شکل ۲-۴. هیستوگرام فراوانی داده‌های کلرید

۳-۲-۴ نیترات (Nitrate)

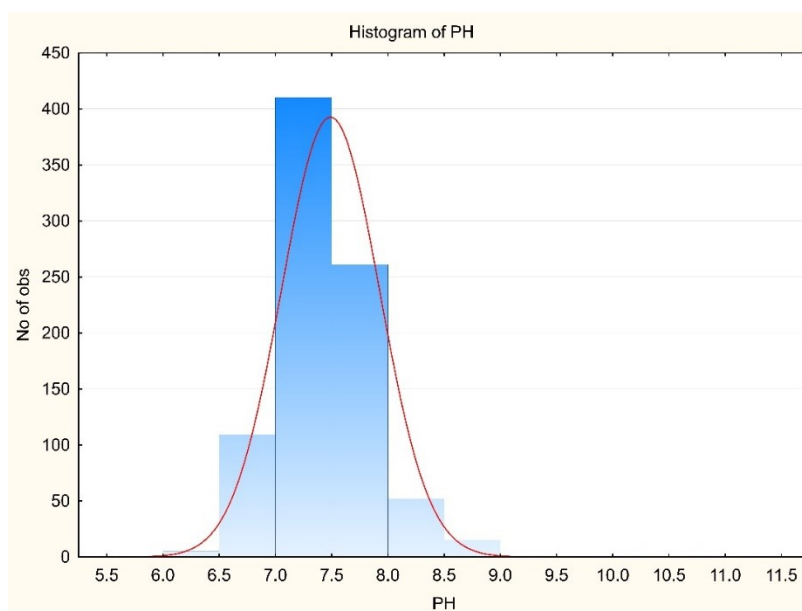
نیترات به طور طبیعی در محیط یافت می شود و از مواد مغذی مهم گیاه است. این ماده در غلظت های مختلف در همه گیاهان وجود دارد و بخشی از چرخه نیتروژن است. نیترات می تواند نتیجه فعالیت های کشاورزی ناشی از دفع فاضلاب و اکسیداسیون فضولات حیوانی و انسانی در سپتیک باشد که با آب های سطحی و زیرزمینی می رسد. غلظت نیترات آب سطحی به دلیل رواناب سطحی جذب کود توسط فیتوپلانکتون و ضد عفونی شدن توسط باکتری ها می تواند به سرعت تغییر کند اما غلظت آب های زیرزمینی عموماً تغییرات نسبتاً کندی را نشان می دهد. همچنین ممکن است آلودگی نیترات در بعضی از آب های زیرزمینی نتیجه شستشوی پوشش گیاهی باشد [۳۸].



شکل ۳-۴. هیستوگرام فراوانی داده های نیترات

۴-۲-۴ پی اچ (PH)

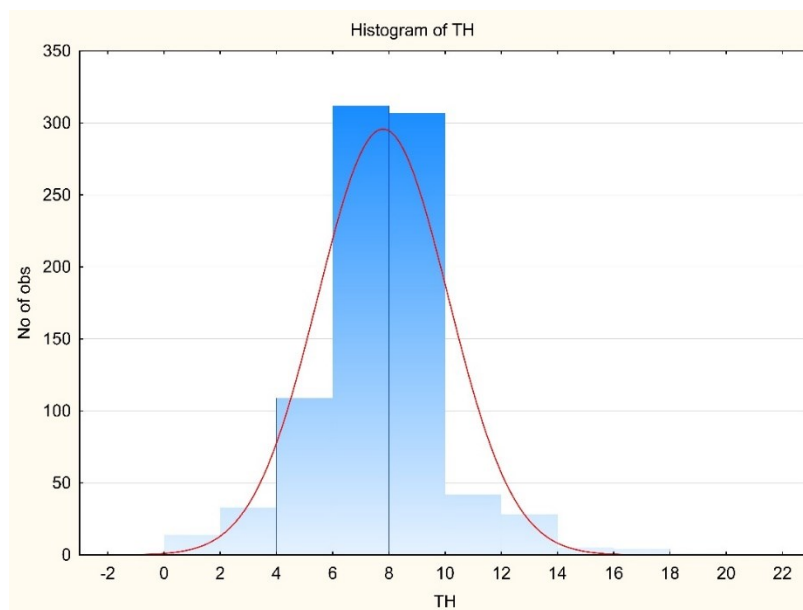
برای PH هیچ دستورالعمل مبتنی بر مقدار مجاز برای آن پیشنهاد نشده است. معمولاً pH تأثیر مستقیمی بر مصرف کنندگان ندارد، اما یکی از مهمترین پارامترهای موثر در کیفیت آب می باشد. باتوجه به استاندارد سازمان بهداشت جهانی مقدار مجاز برای PH ۶/۵ تا ۸/۵ بوده که عدد ۷ نشان دهنده خنثی بودن محلول می باشد. با افزایش عدد PH در محلول خاصیت بازی محلول افزایش می یابد و با کاهش آن بر خاصیت اسیدی محلول افزون می گردد [۳۸].



شکل ۴-۴. هیستوگرام فراوانی داده های PH

۵-۲-۴ سختی کل (TH)

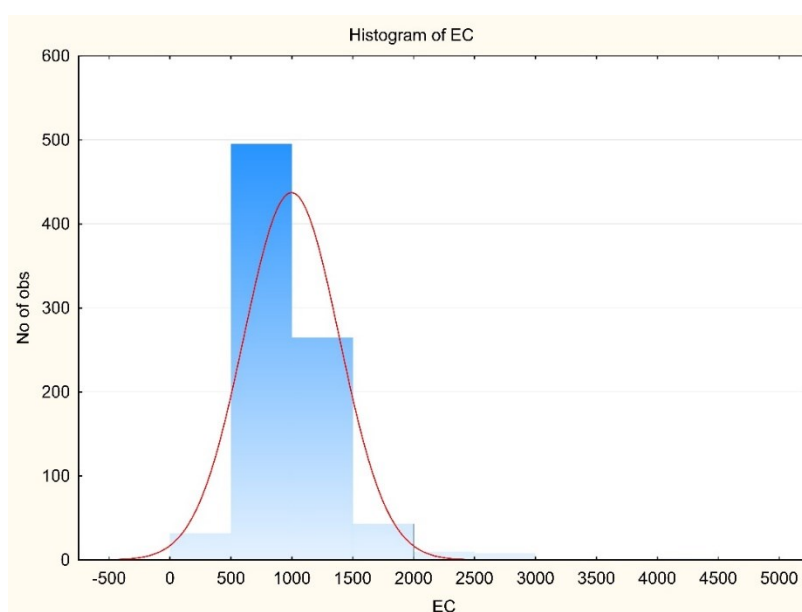
سختی ناشی از کلسیم و منیزیم معمولاً با رسوب کف صابون و پایداری در استفاده زیاد از صابون برای تمیزی به وجود می‌آید. پذیرش عمومی درجه سختی آب ممکن است به طور قابل توجهی از یک جامعه به جامعه دیگر متفاوت باشد. یون کلسیم متناسب با آنیون مرتبط در محدوده ۱۰۰-۳۰۰ میلی‌گرم در لیتر بوده که احتمالاً این بازه برای منیزیم کمتر می‌باشد. در برخی موارد، سختی آب بیش از ۵۰۰ میلی‌گرم در لیتر برای مصرف کنندگان قابل تحمل است. آب با سختی بالاتر از مقدار ۲۰۰ میلی‌گرم در لیتر بسته به فعل و انفعال عوامل دیگر مانند PH و قلیایی بودن، ممکن است باعث رسوب در کارهای تصفیه، سیستم توزیع و لوله گذاری و مخازن داخل ساختمان شود [۳۸].



شکل ۴-۵. هیستوگرام فراوانی داده‌های سختی کل

۶-۲-۴ نسبت جذب سدیم (SAR)

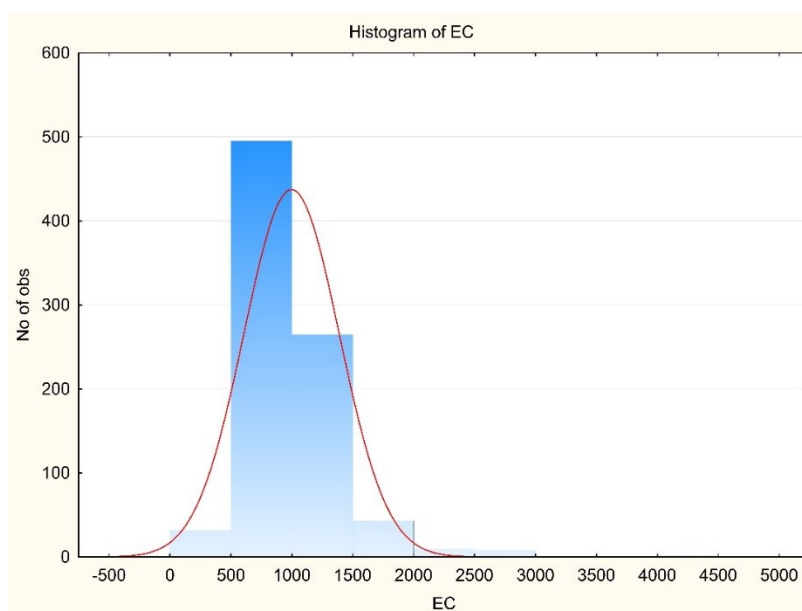
نسبت جذب سدیم به عنوان شاخصی موثر در تعیین سدیم می باشد که با افزایش غلظت یون مورد نظر در آب، مقدار آن افزایش می یابد. مقدار مجاز برای نسبت جذب سدیم متناسب با TDS از ۲ تا ۱۰۰ متغیر می باشد. استاندارد این شاخص برای گونه های حساس بین ۸ تا ۱۷ در نظر گرفته شده است [۳۹].



شکل ۴-۶. هیستوگرام فراوانی داده های SAR

۷-۲-۴ هدایت الکتریکی (EC)

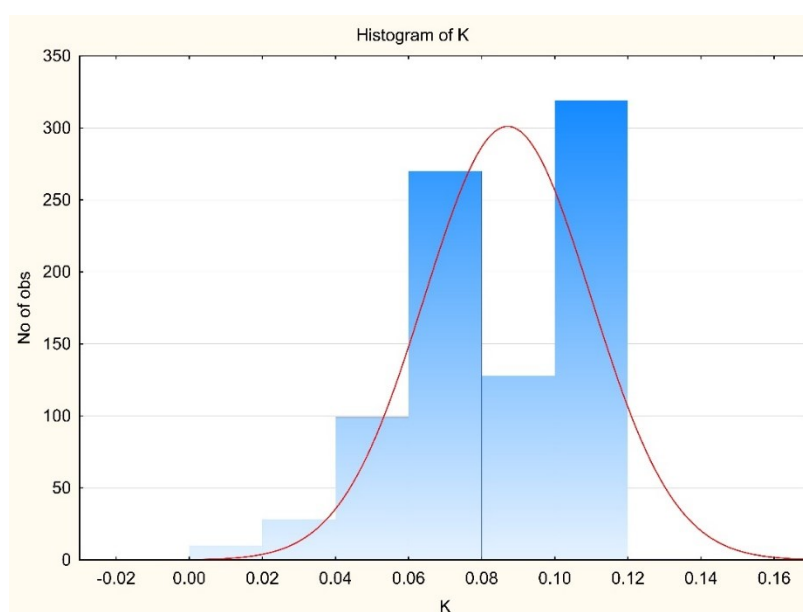
قابلیت رسانایی، اندازه گیری توانایی آب در عبور جریان الکتریکی بوده که تحت تأثیر مواد جامد محلول غیر آلی مانند آنیون‌های کلرید، نیترات، سولفات و فسفات (یون‌هایی که بار منفی دارند) یا کاتیون‌های سدیم، منیزیم، کلسیم، آهن و آلومینیوم (یون‌هایی که دارای بار مثبت هستند) می‌باشد. این قابلیت تحت تأثیر دما قرار دارد: هرچه آب گرم‌تر باشد، رسانایی نیز بالاتر است. به همین دلیل اندازه گیری رسانایی در دمای ۲۵ درجه انجام می‌شود [۴۰].



شکل ۷-۴. هیستوگرام فراوانی داده‌های EC

۸-۲-۴ پتاسیم (Potassium)

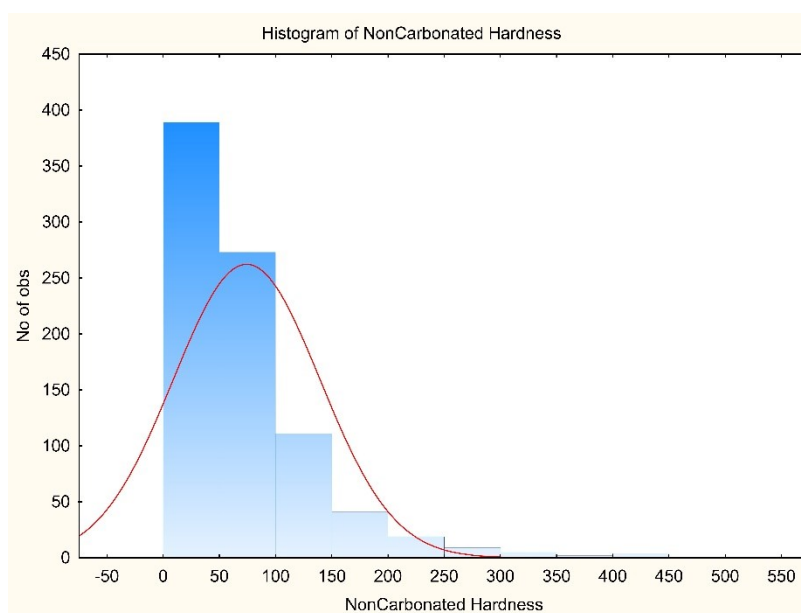
پتاسیم یک عنصر اساسی در انسان است و به ندرت در آب آشامیدنی در سطحی یافت می شود که بتواند برای سلامتی انسان نگران کننده باشد. نیاز روزانه توصیه شده بیشتر از ۳۰۰۰ میلی گرم است که به طور گسترده در محیط از جمله تمام آب های طبیعی وجود دارد. ممکن است پتاسیم برخی از تاثیرات منفی را در افراد حساس ایجاد کند با این حال مصرف پتاسیم از آب آشامیدنی کاملاً کمتر از سطحی است که عوارض جانبی بر روی سلامتی داشته باشد [۳۸].



شکل ۴-۸. هیستوگرام فراوانی داده های K

۹-۲-۴ سختی موقت و دائم آب

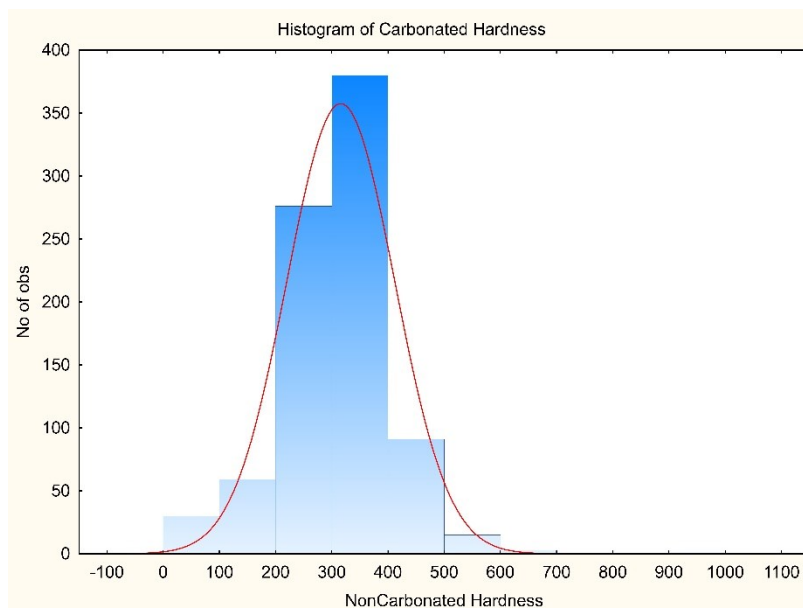
گاهی اوقات سختی آب به سختی کلسیم و سختی منیزیم تفکیک می‌شود. این عمل با کسر سختی کلسیم از سختی کل برای محاسبه سختی منیزیم انجام می‌شود. سختی کلسیم را می‌توان از غلظت کلسیم یا مستقیماً توسط تیتراسیون اتیلن‌دی‌آمین تترا استیک اسید (EDTA) با استفاده از نشانگر مورکسید در نمونه‌ای که منیزیم در اثر pH بالا از آن خارج شده است، اندازه‌گیری کرد. نشانگر مورکسید برای پایداری در pH بالا در رسوب منیزیم لازم است اما عملکرد آن همانند Eriochrome Black T^۲ می‌باشد. همچنین سختی می‌تواند به سختی کربناته و غیر کربناته تقسیم شود [۳۹].



شکل ۹-۴. هیستوگرام فراوانی داده‌های سختی دائم

^۱Ethylenediaminetetraacetic Acid

^۲ شاخصی است که در فرآیند تعیین سختی آب قابل استفاده می‌باشد.

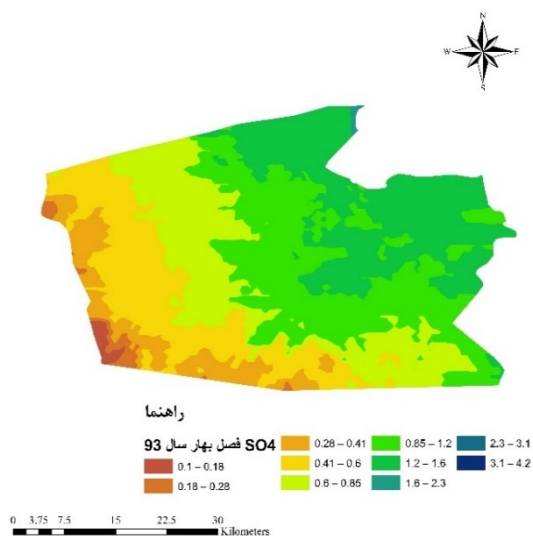


شکل ۴-۱۰. هیستوگرام فراوانی داده‌های سختی موقت

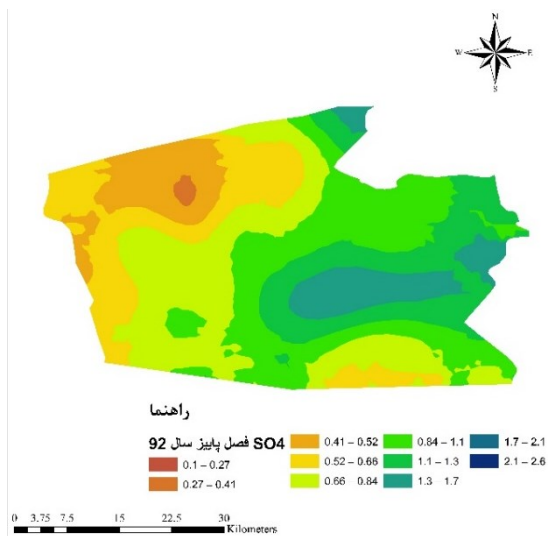
۴-۳ پهنه بندی در GIS و استخراج داده‌ها

باتوجه به اهمیت تعداد پارامترهای مورد نیاز در سطح منطقه مورد مطالعه برای بررسی بهتر روند تغییرات نیاز به اطلاعات بیشتری می‌باشد. با استفاده از درونیابی مقادیر جمع آوری شده از چاه‌های مورد نظر روند تغییرات پارامترها در شش سال، بدست آمده است. بدین صورت که به ازای هر سال آبی دو سری داده اندازه گیری شده که هر یک از داده‌ها در محدوده منطقه مورد مطالعه پهنه‌بندی شده می‌شود. برای انجام این عمل، از نرم افزار ArcGIS جهت ورود اطلاعات چاه‌ها و لایه‌های اطلاعاتی به ازای هر پارامتر استفاده شده است.

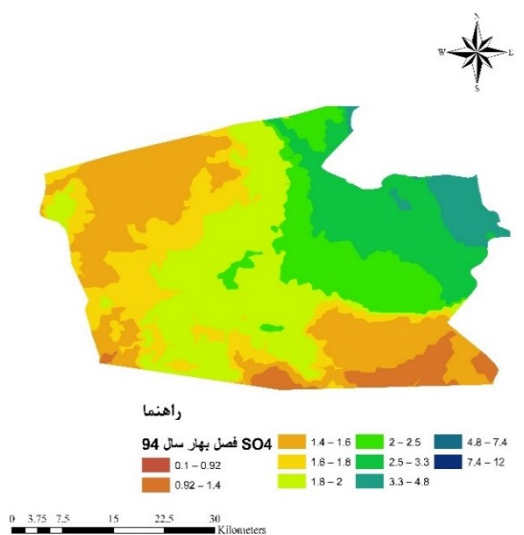
در این تحقیق از روش kriging به دلیل کمترین مقدار خطای جذر میانگین مربعات (RMSE) و بیشترین مقدار ضریب تشخیص (R^2) نسبت به سایر روش‌ها برای درونیابی استفاده شده است. جهت انجام درونیابی در منطقه مورد نظر و ایجاد پهنه بندی مناسب و مجزا برای هر پارامتر، مرز درونیابی به محدوده‌ای نزدیک به چاه‌های مورد نظر محدود شده است [۴۲]. نتایج حاصل از پهنه‌بندی برای پارامترهای مورد نظر در شکل ۴-۱۱ تا ۴-۱۵۹ آورده شده است.



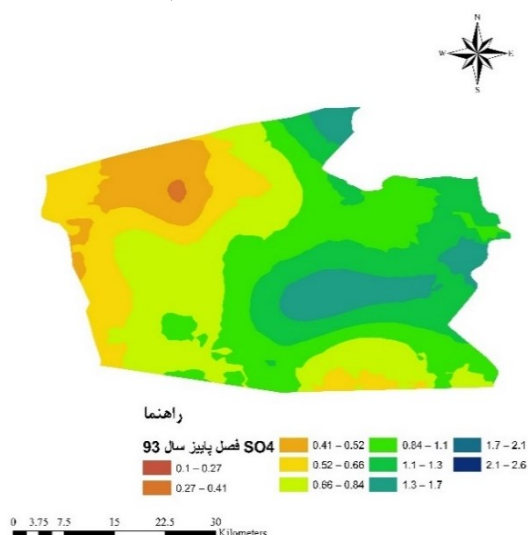
شکل ۴-۱۲. سولفات فصل بهار ۹۳



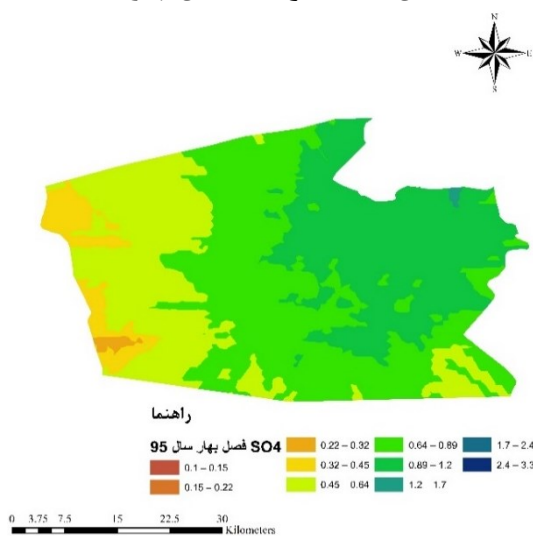
شکل ۴-۱۱. سولفات فصل پاییز ۹۲



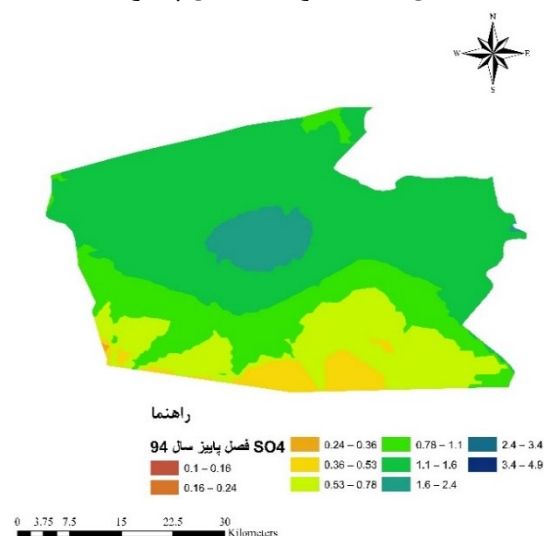
شکل ۴-۱۴. سولفات فصل بهار ۹۴



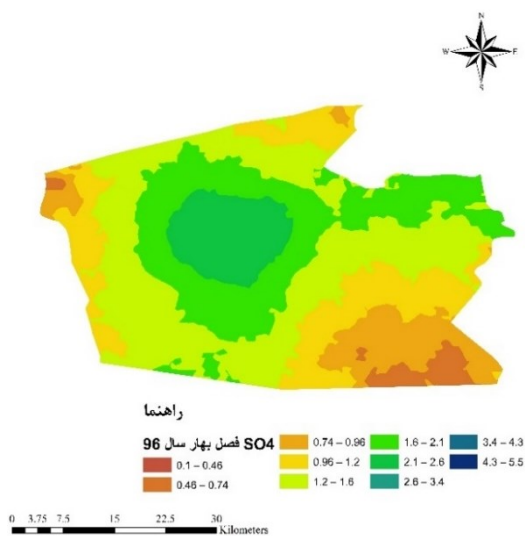
شکل ۴-۱۳. سولفات فصل پاییز ۹۳



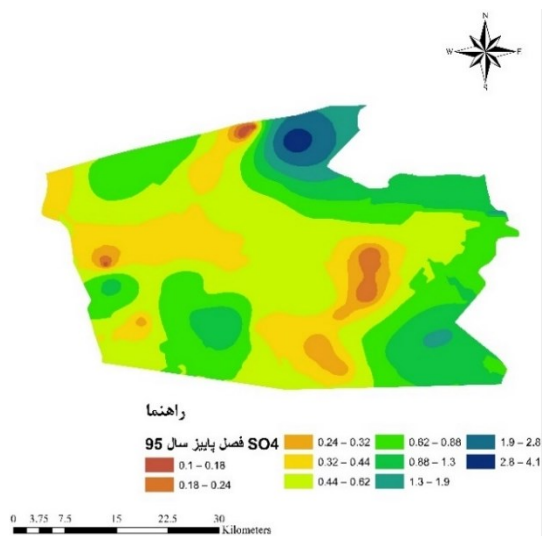
شکل ۴-۱۶. سولفات فصل بهار ۹۵



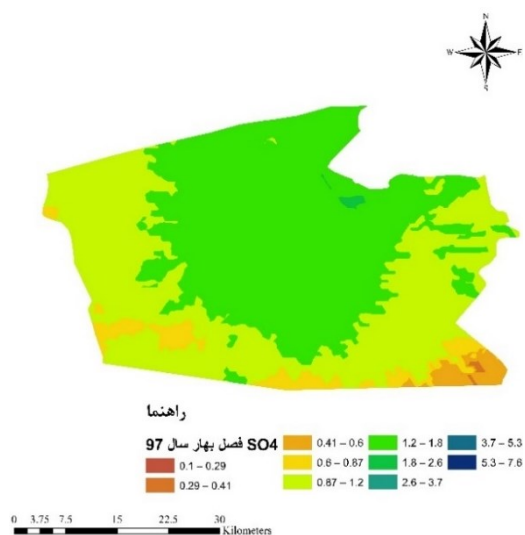
شکل ۴-۱۵. سولفات فصل پاییز ۹۴



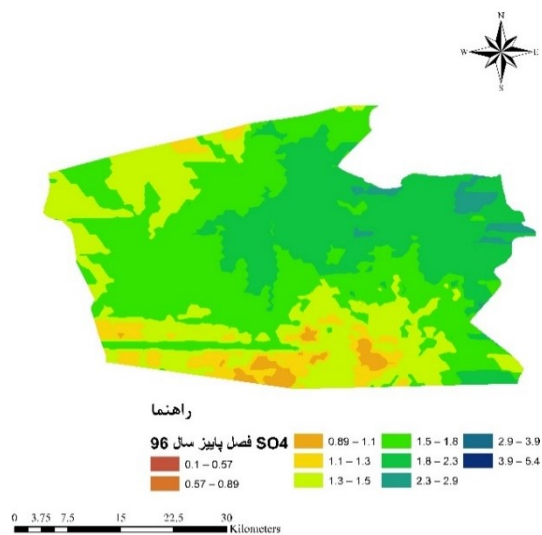
شکل ۴-۱۸. سولفات فصل بهار ۹۶



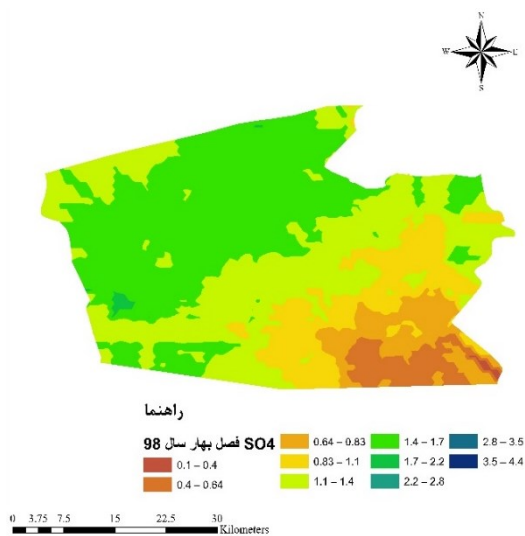
شکل ۴-۱۷. سولفات فصل پاییز ۹۵



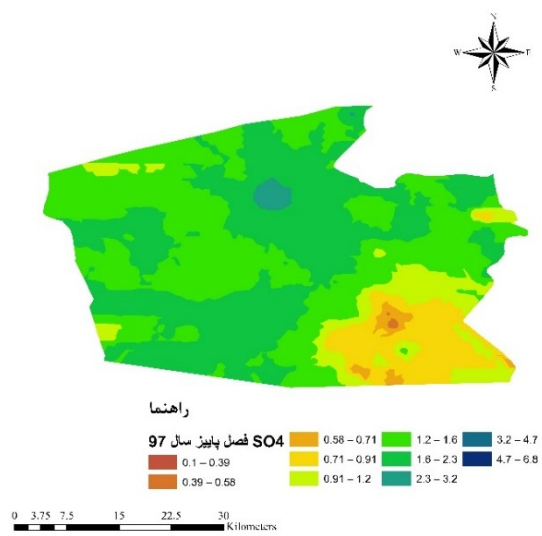
شکل ۴-۲۰. سولفات فصل بهار ۹۷



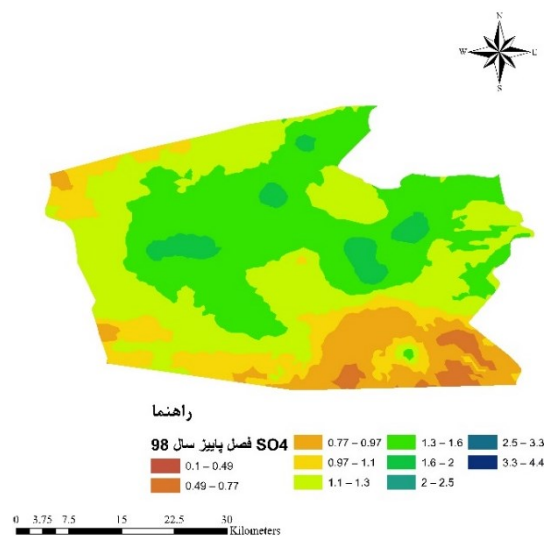
شکل ۴-۱۹. سولفات فصل پاییز ۹۶



شکل ۴-۲۲. سولفات فصل بهار ۹۸

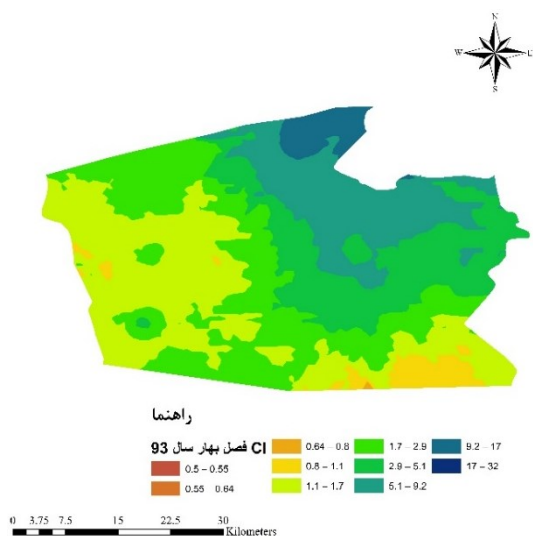


شکل ۴-۲۱. سولفات فصل پاییز ۹۷

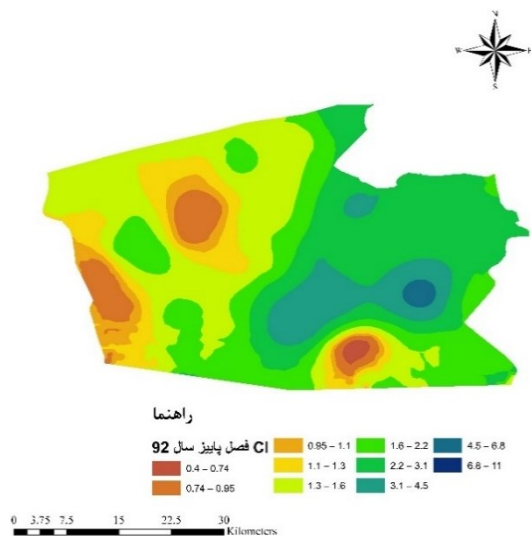


شکل ۴-۲۳. سولفات فصل پاییز ۹۸

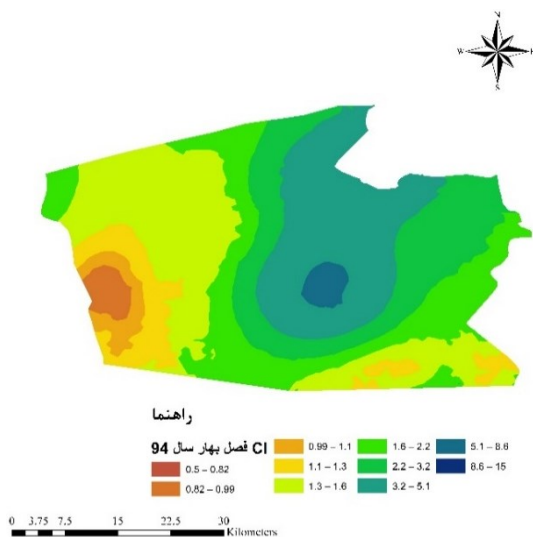
همانگونه که در شکل ۴-۱۱ تا ۴-۲۳ مشخص است، تغییرات در پهنه‌بندی سولفات برای پاییز ۹۲ و ۹۳ تقریباً مشابه است. این روند برای بهار ۹۳ و ۹۴ تقریباً بدین گونه بوده، بطوری که در قسمت شرقی الگوی تغییرات مشابه و در سایر بخش‌ها متفاوت می‌باشد. برای سایر سال‌ها تغییرات متفاوت بوده و الگوی خاصی نمی‌توان در نظر گرفت.



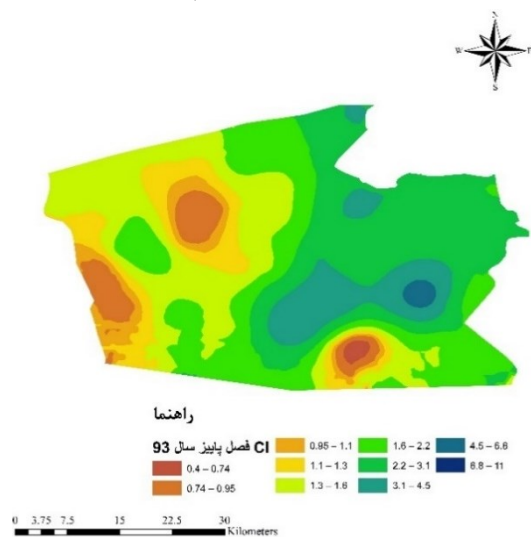
شکل ۴-۲۵. کمر فصل بهار ۹۳



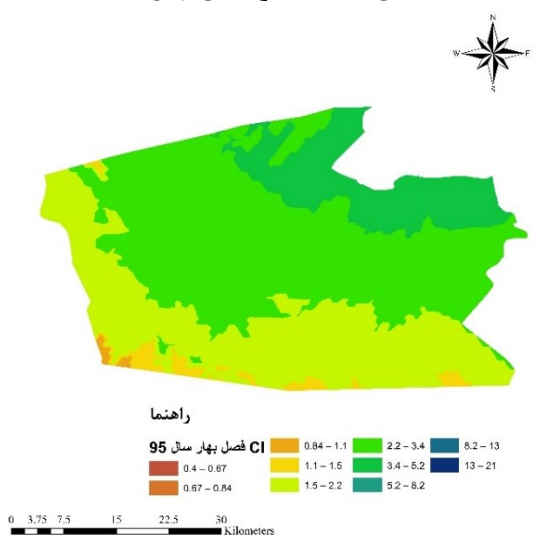
شکل ۴-۲۴. کمر فصل پاییز ۹۲



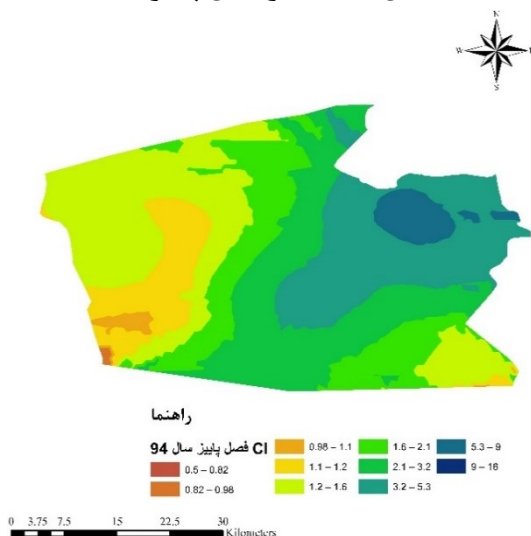
شکل ۴-۲۷. کمر فصل بهار ۹۴



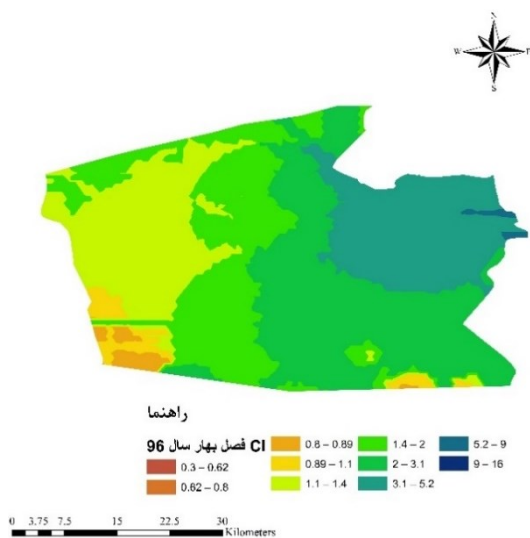
شکل ۴-۲۶. کمر فصل پاییز ۹۳



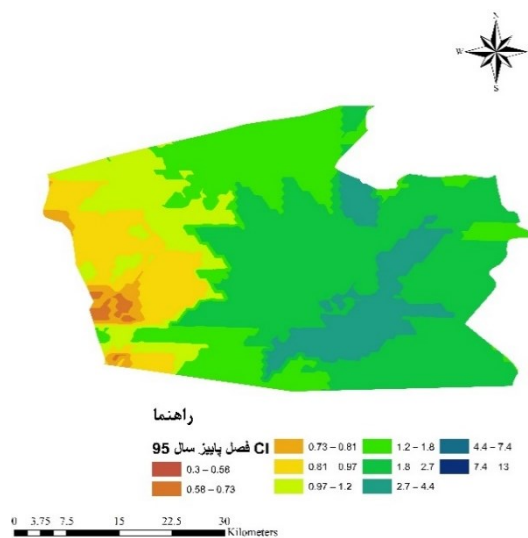
شکل ۴-۲۹. کمر فصل بهار ۹۵



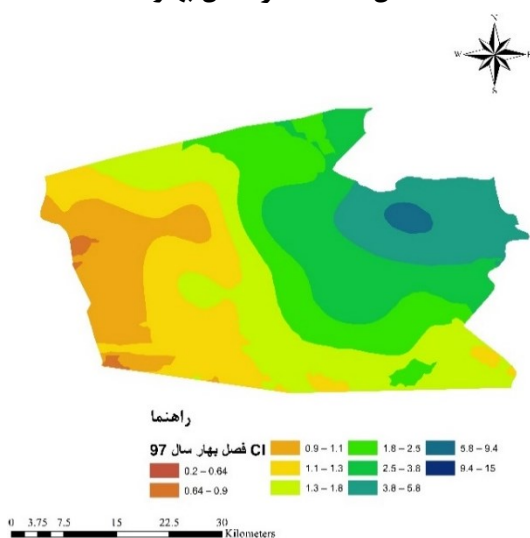
شکل ۴-۲۸. کمر فصل پاییز ۹۴



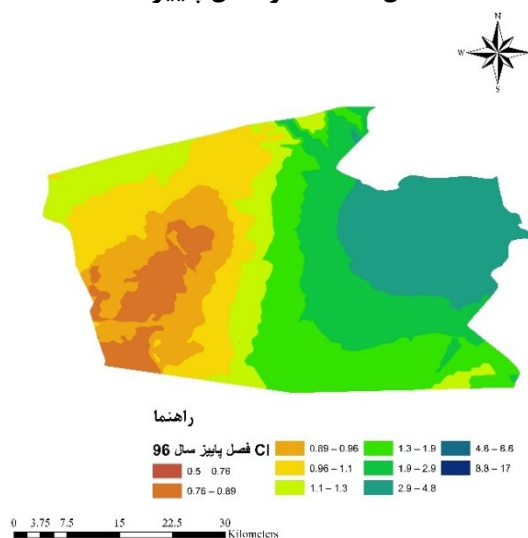
شکل ۴-۳۱. کمر فصل بهار ۹۶



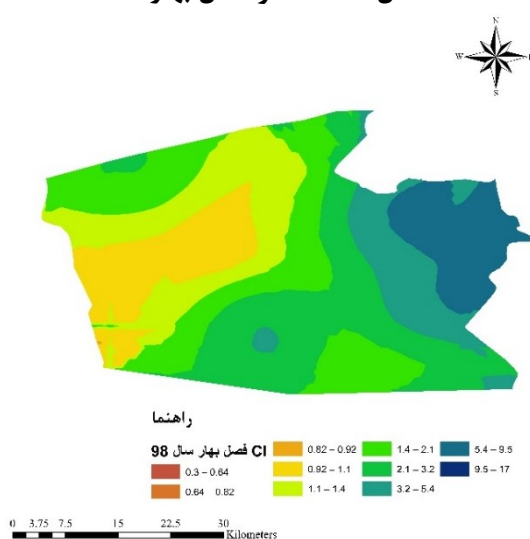
شکل ۴-۳۰. کمر فصل پاییز ۹۵



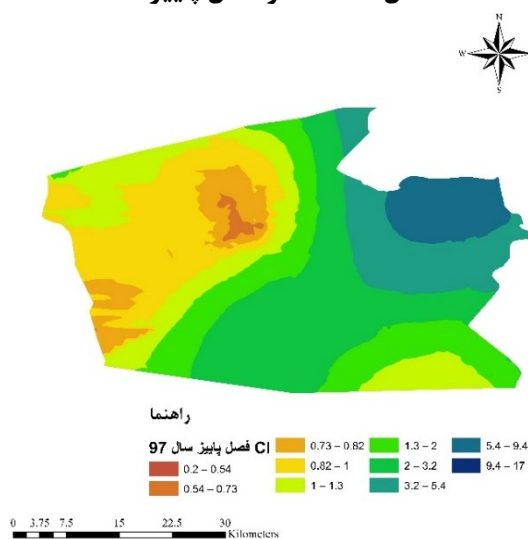
شکل ۴-۳۳. کمر فصل بهار ۹۷



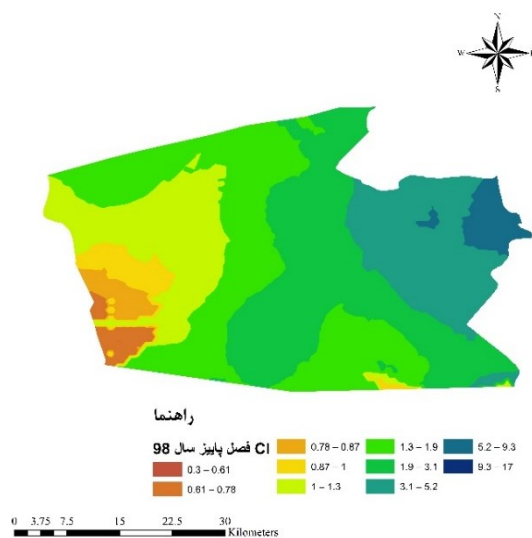
شکل ۴-۳۲. کمر فصل پاییز ۹۶



شکل ۴-۳۵. کمر فصل بهار ۹۸



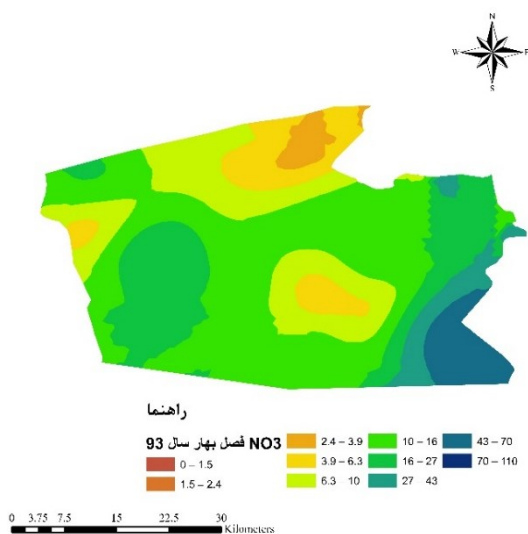
شکل ۴-۳۴. کمر فصل پاییز ۹۷



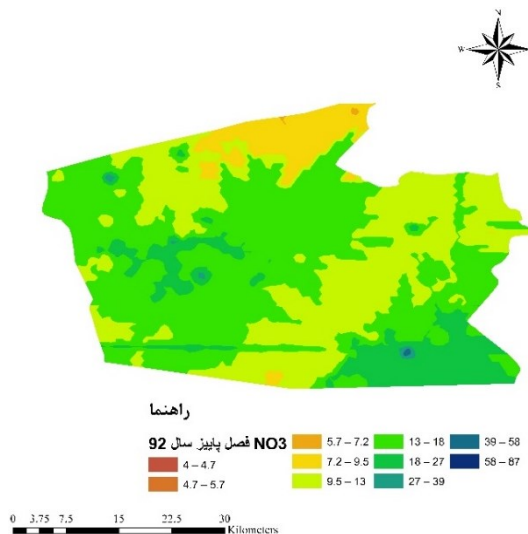
شکل ۴-۳۶. کلر فصل پاییز ۹۸

شکل ۴-۲۴ تا ۴-۳۶ نشان دهنده پهنه بندی کلر است که باتوجه به آن در برخی فصلها نتایج

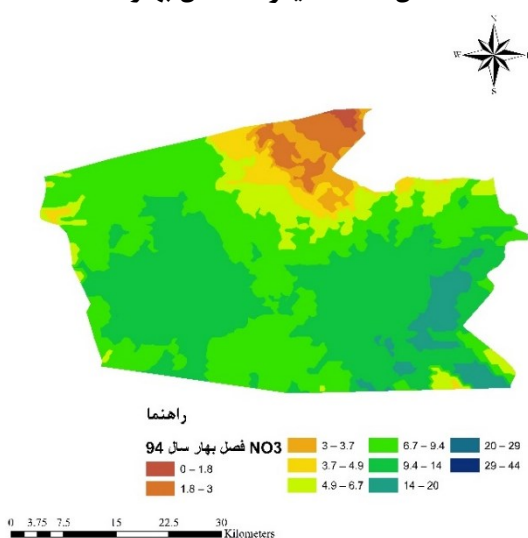
مشابه یکدیگر می باشد.



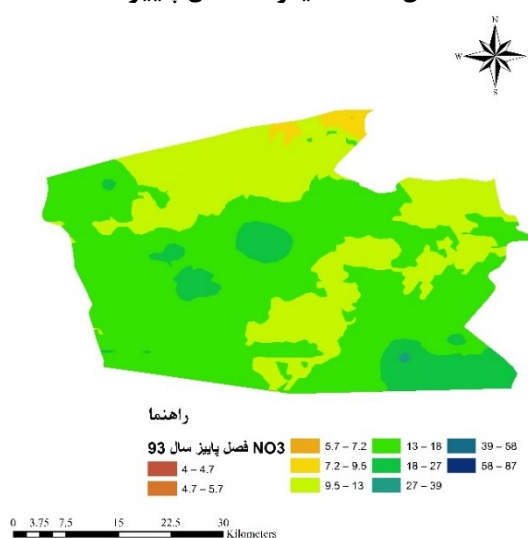
شکل ۴-۳۸. نیترات فصل بهار ۹۳



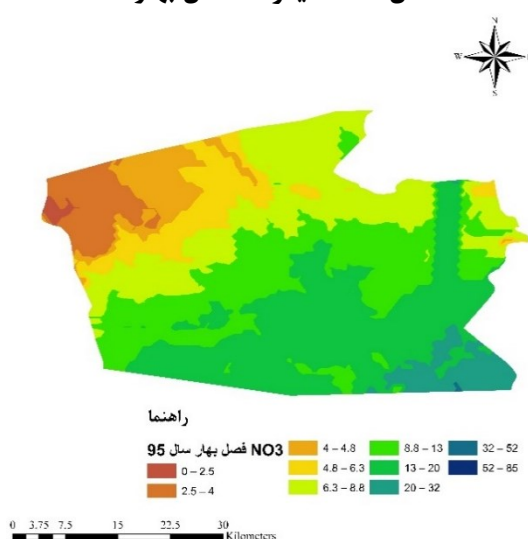
شکل ۴-۳۷. نیترات فصل پاییز ۹۲



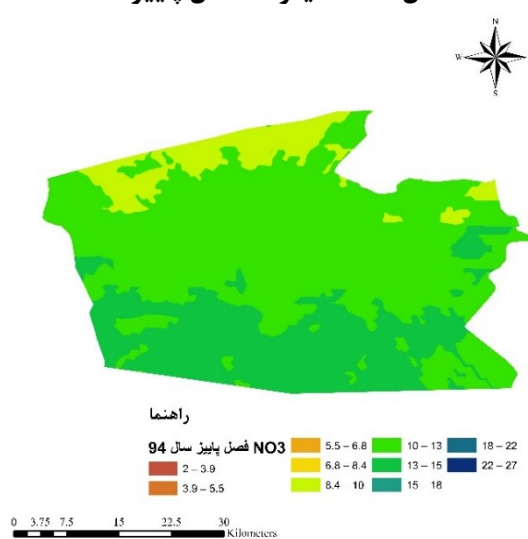
شکل ۴-۴۰. نیترات فصل بهار ۹۴



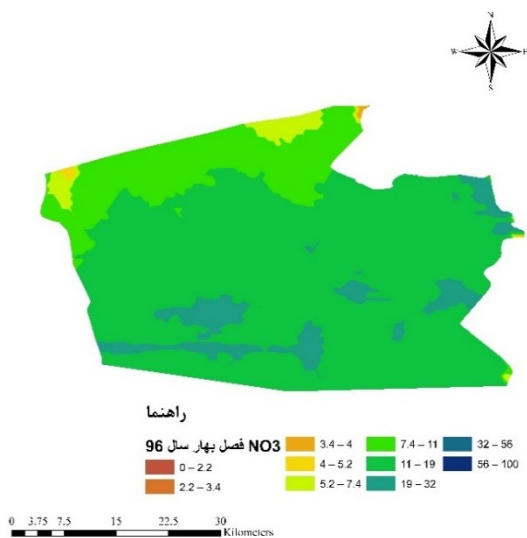
شکل ۴-۳۹. نیترات فصل پاییز ۹۳



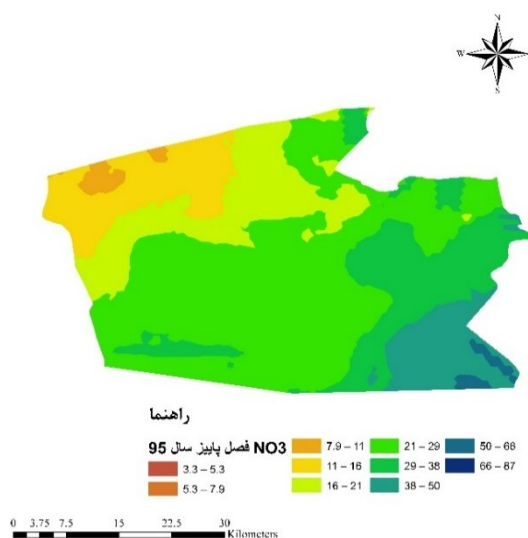
شکل ۴-۴۲. نیترات فصل بهار ۹۵



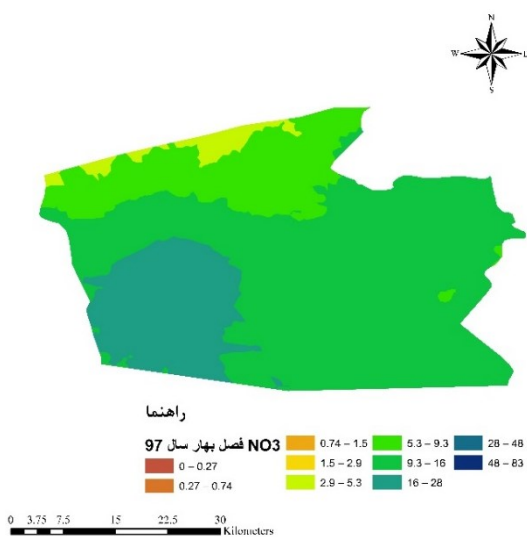
شکل ۴-۴۱. نیترات فصل پاییز ۹۴



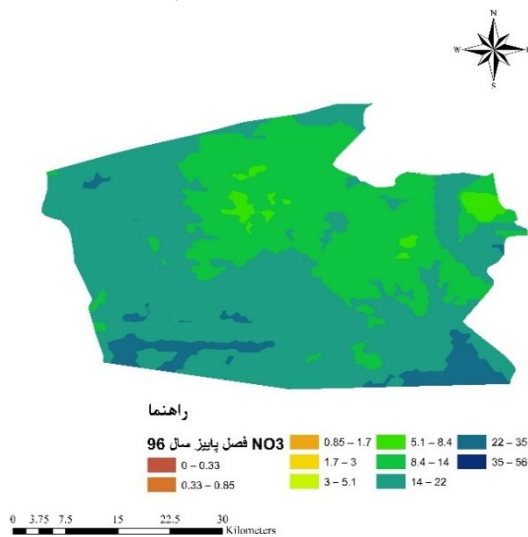
شکل ۴-۴۴. نیترات فصل بهار ۹۶



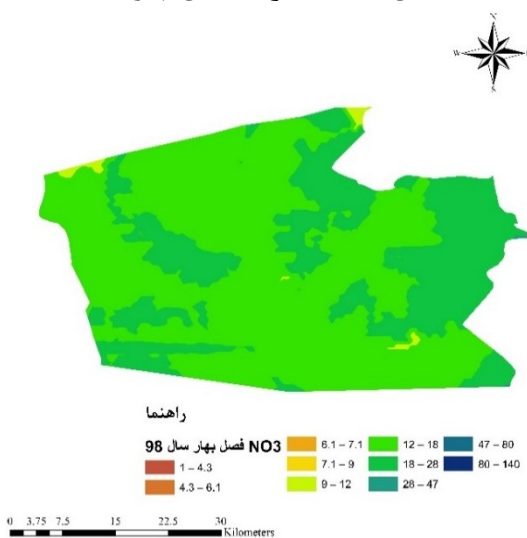
شکل ۴-۴۳. نیترات فصل پاییز ۹۵



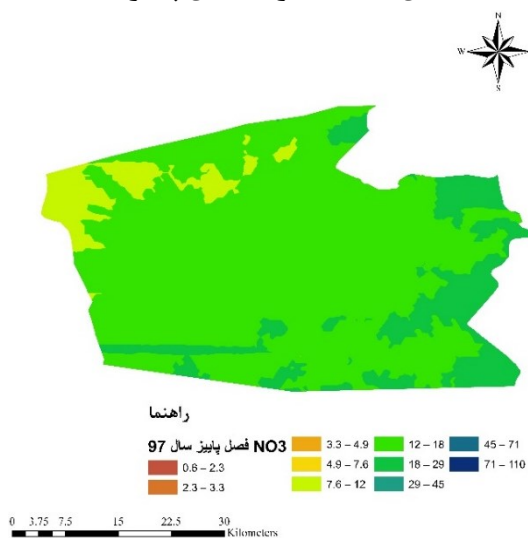
شکل ۴-۴۶. نیترات فصل بهار ۹۷



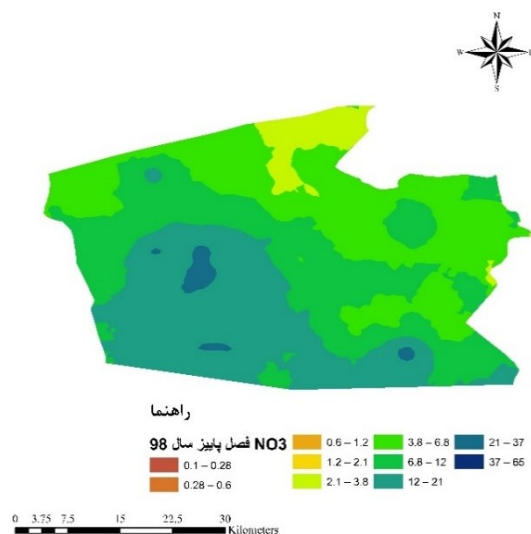
شکل ۴-۴۵. نیترات فصل پاییز ۹۶



شکل ۴-۴۸. نیترات فصل بهار ۹۸

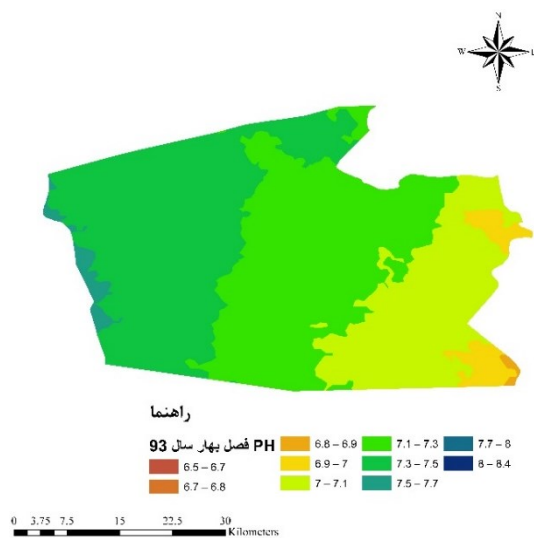


شکل ۴-۴۷. نیترات فصل پاییز ۹۷

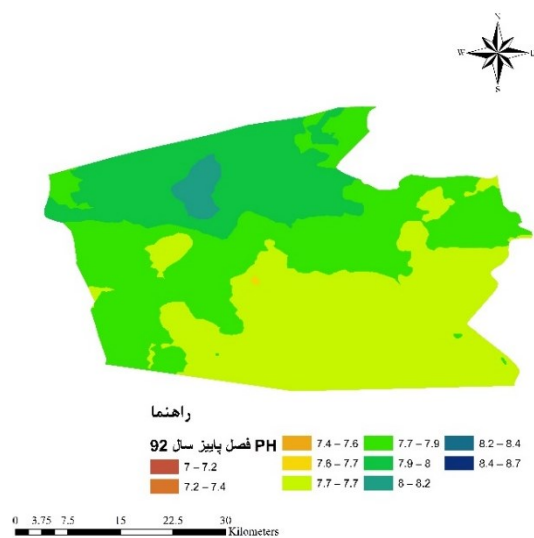


شکل ۴-۴۹. نیترات فصل پاییز ۹۸

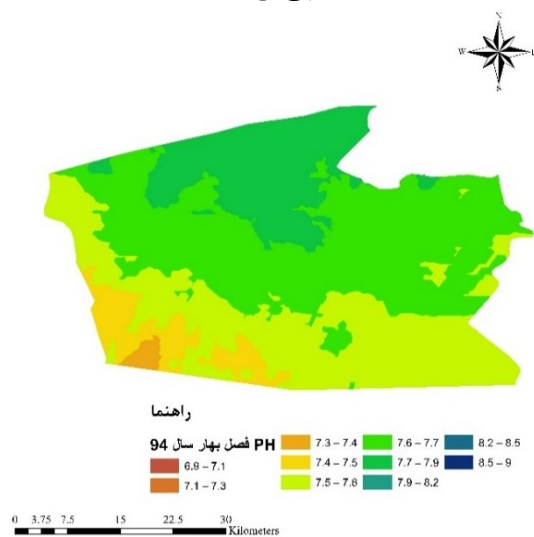
برای الگوی پراکندگی نیترات در شکل ۴-۳۷ تا ۴-۴۹ نشان می‌دهد که تغییرات در تمام سال‌ها به جز پاییز ۹۲ و ۹۵، بهار سال ۹۳، ۹۴ و ۹۵ الگوی مشابه دارد. همچنین در گذر از فصل پاییز به فصل بهار به دلیل کوددهی و بارش باران شاهد افزایش غلظت نیترات هستیم.



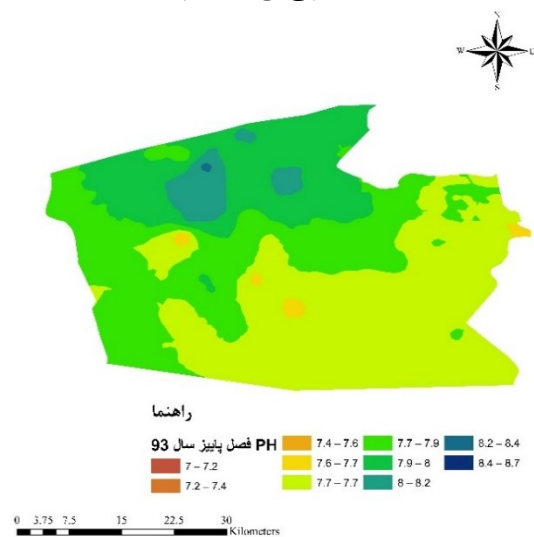
شکل ۴-۵۱. پی اچ فصل بهار ۹۳



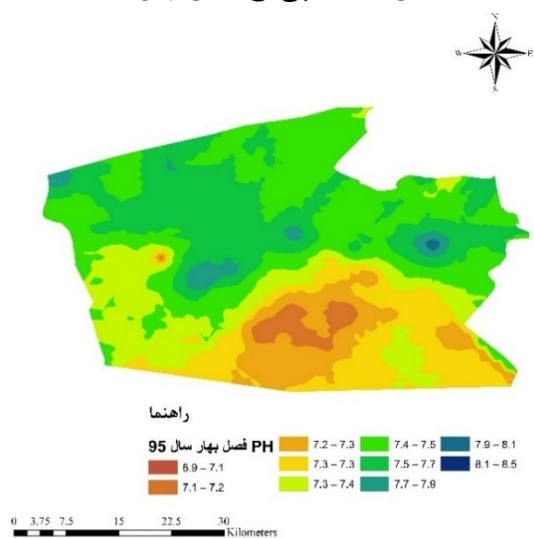
شکل ۴-۵۰. پی اچ فصل پاییز ۹۲



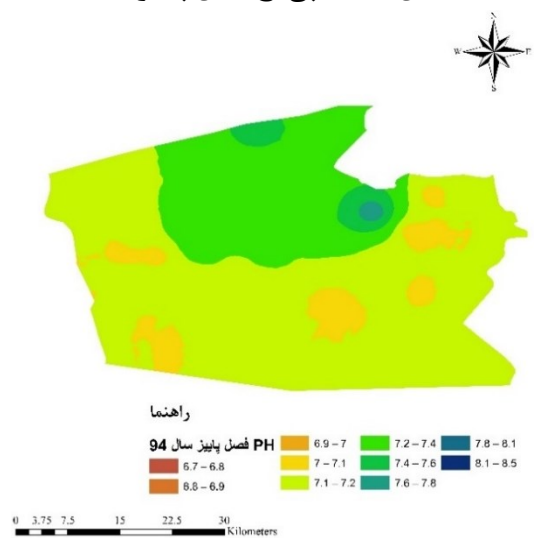
شکل ۴-۵۳. پی اچ فصل بهار ۹۴



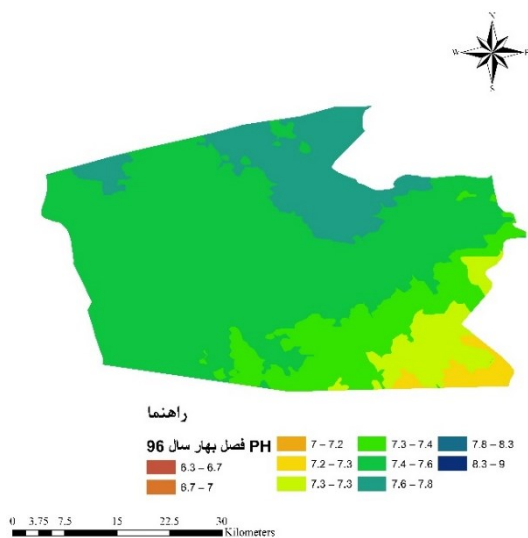
شکل ۴-۵۲. پی اچ فصل پاییز ۹۳



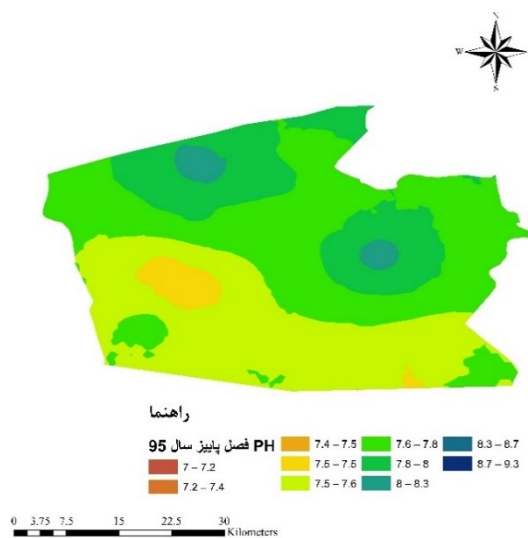
شکل ۴-۵۵. پی اچ فصل بهار ۹۵



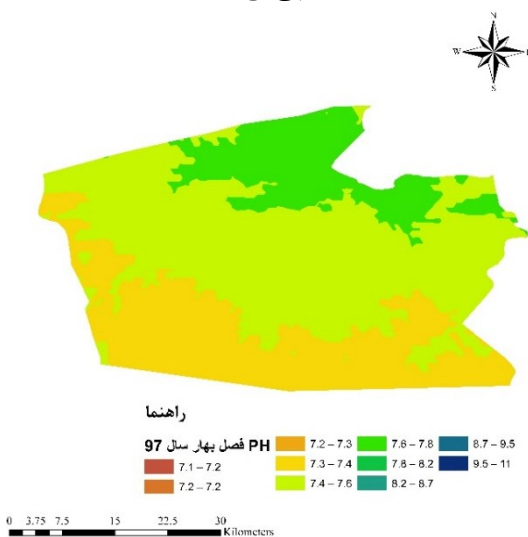
شکل ۴-۵۴. پی اچ فصل پاییز ۹۴



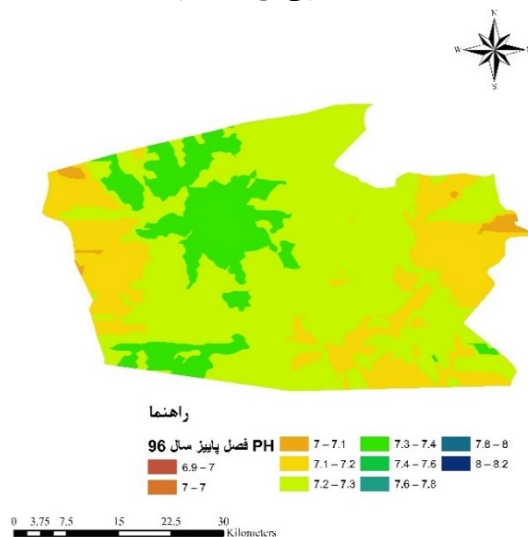
شکل ۴-۵۷. پی اچ فصل بهار ۹۶



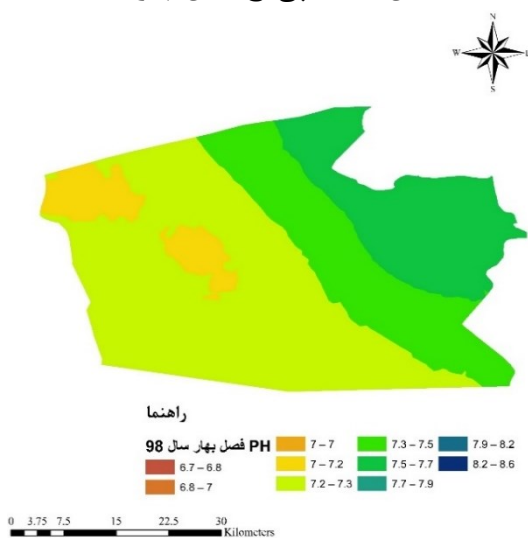
شکل ۴-۵۶. پی اچ فصل پاییز ۹۵



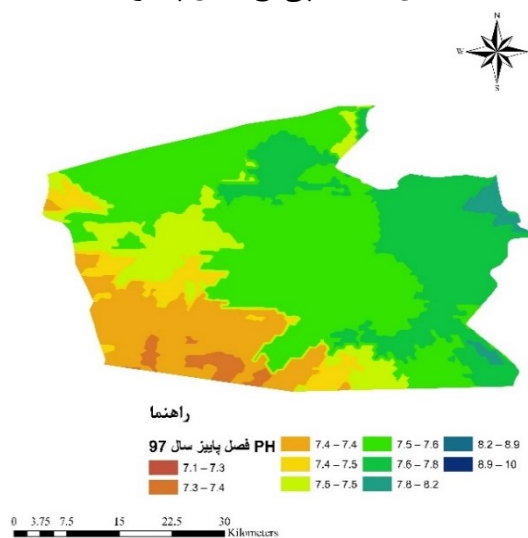
شکل ۴-۵۹. پی اچ فصل بهار ۹۷



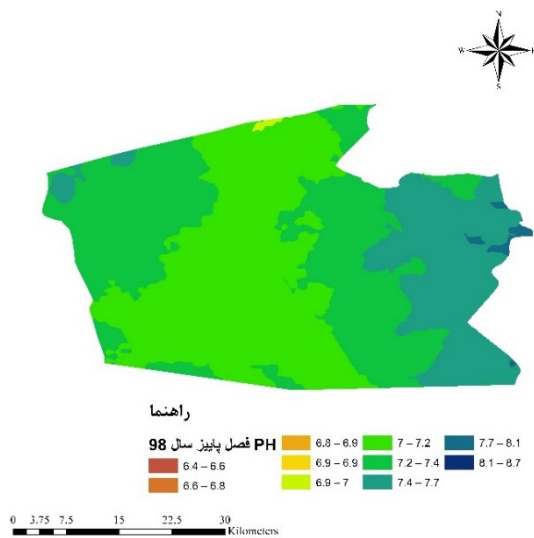
شکل ۴-۵۸. پی اچ فصل پاییز ۹۶



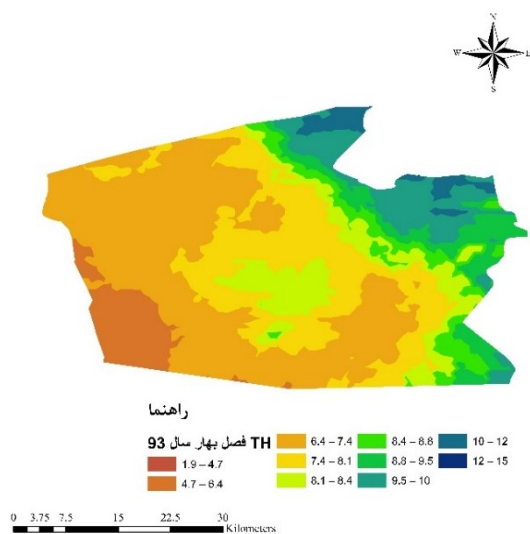
شکل ۴-۶۱. پی اچ فصل بهار ۹۸



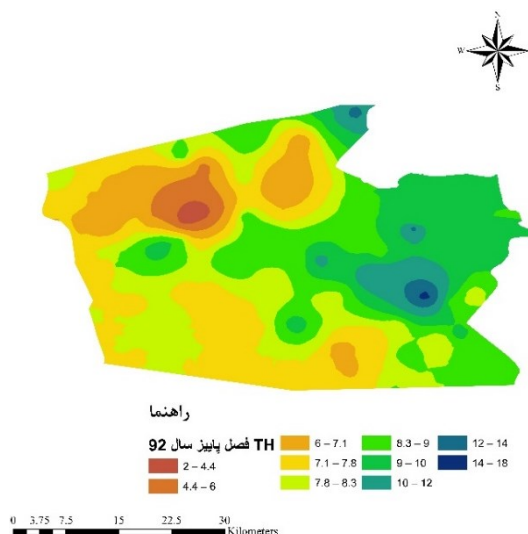
شکل ۴-۶۰. پی اچ فصل پاییز ۹۷



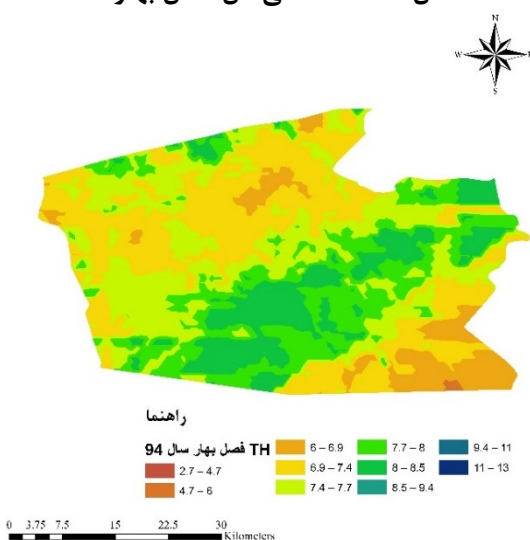
شکل ۴-۶۲. پی اچ فصل پاییز ۹۸



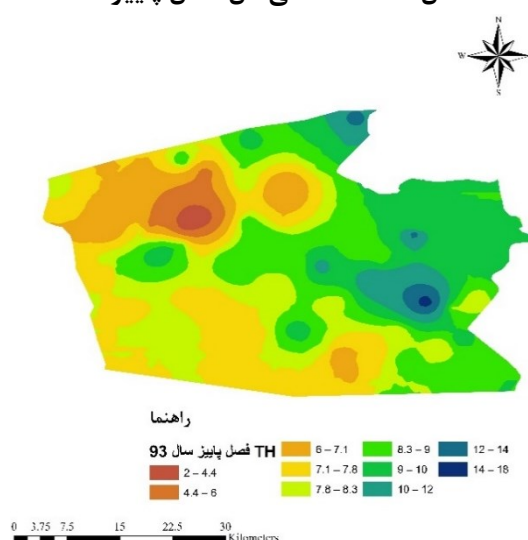
شکل ۴-۶۴. سختی کل فصل بهار ۹۳



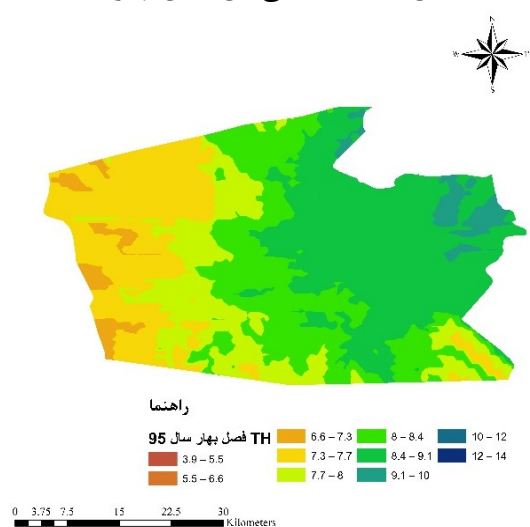
شکل ۴-۶۳. سختی کل فصل پاییز ۹۲



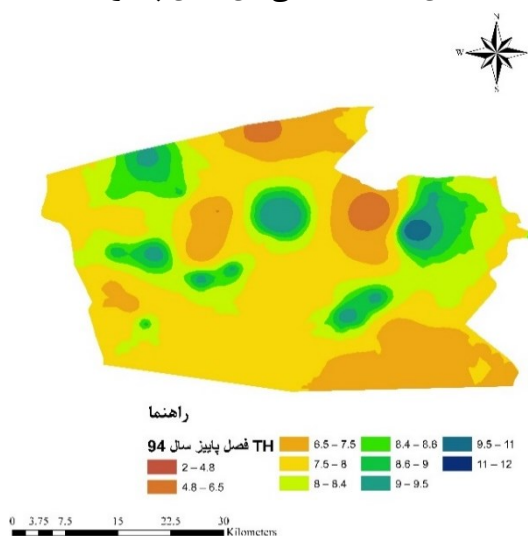
شکل ۴-۶۶. سختی کل فصل بهار ۹۴



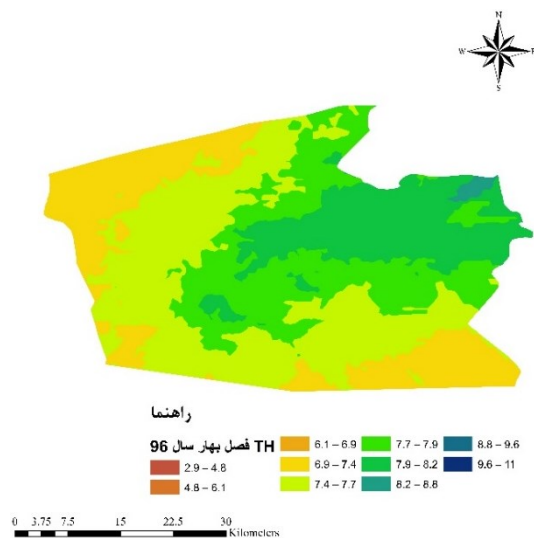
شکل ۴-۶۵. سختی کل فصل پاییز ۹۳



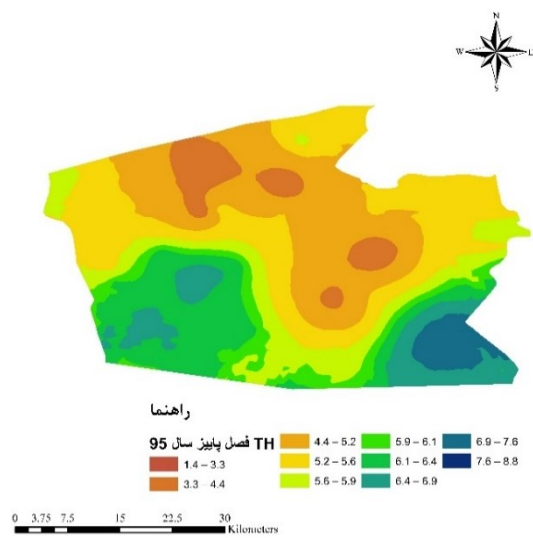
شکل ۴-۶۸. سختی کل فصل بهار ۹۵



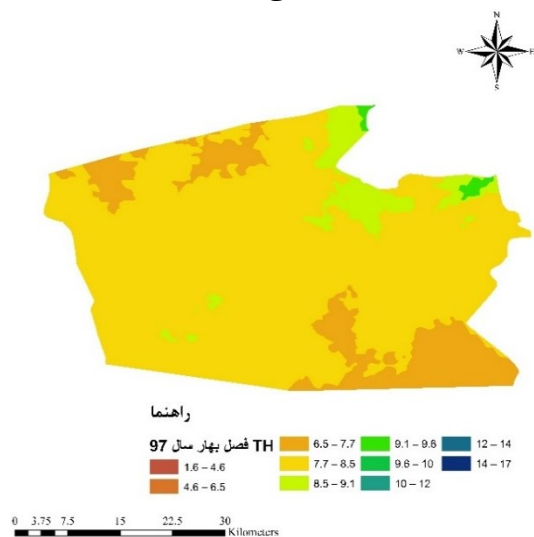
شکل ۴-۶۷. سختی کل فصل پاییز ۹۴



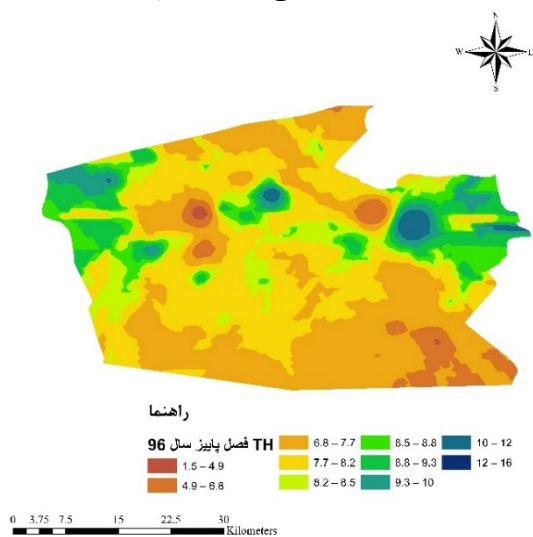
شکل ۴-۷۰. سختی کل فصل بهار ۹۶



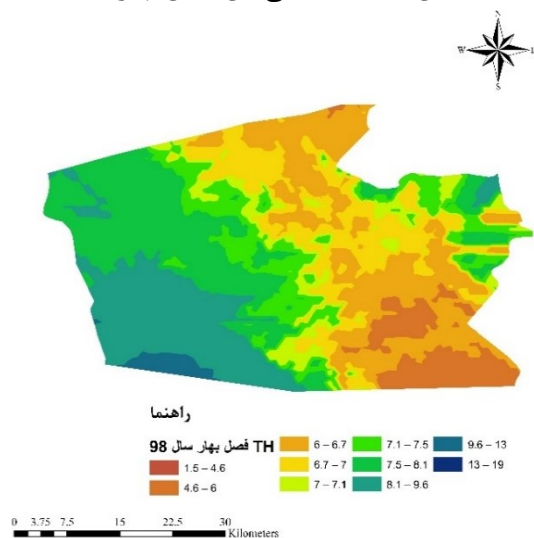
شکل ۴-۹۵. سختی کل فصل پاییز ۹۵



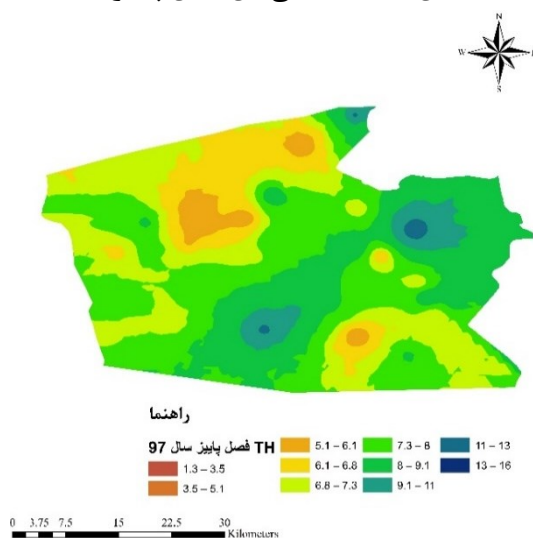
شکل ۴-۷۲. سختی کل فصل بهار ۹۷



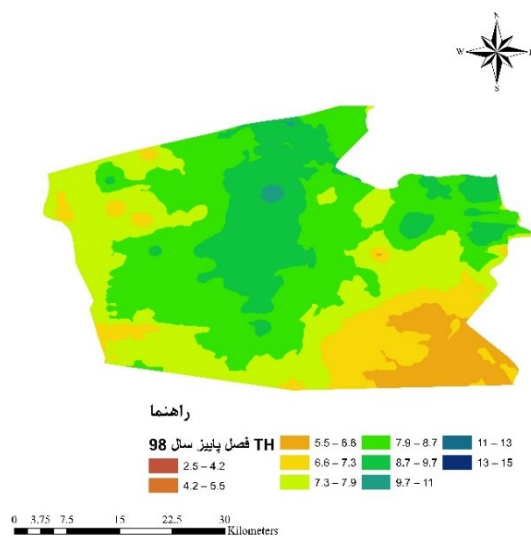
شکل ۴-۷۱. سختی کل فصل پاییز ۹۶



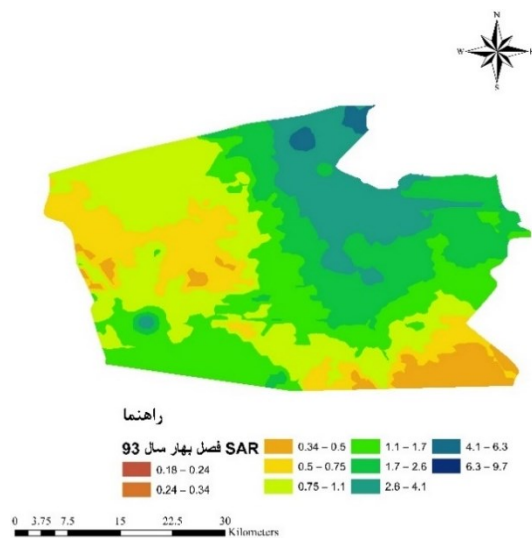
شکل ۴-۷۴. سختی کل فصل بهار ۹۸



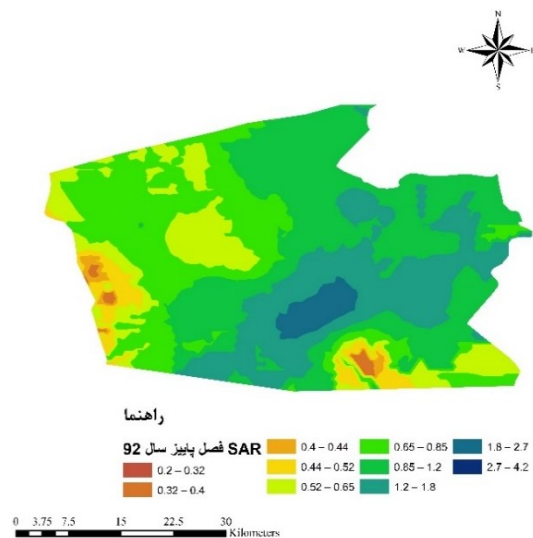
شکل ۴-۷۳. سختی کل فصل پاییز ۹۷



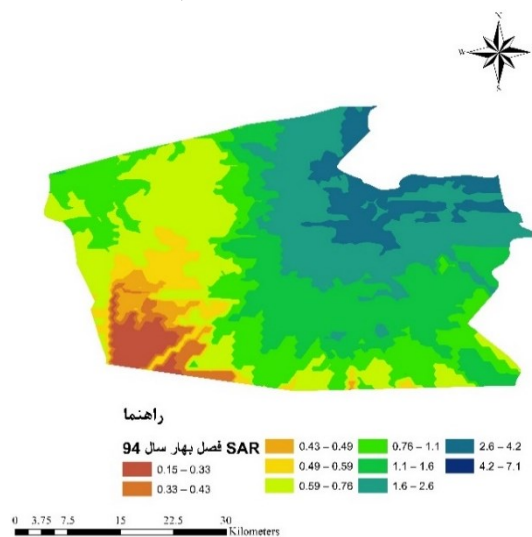
شکل ۷۵-۴. سختی کل فصل پاییز ۹۸



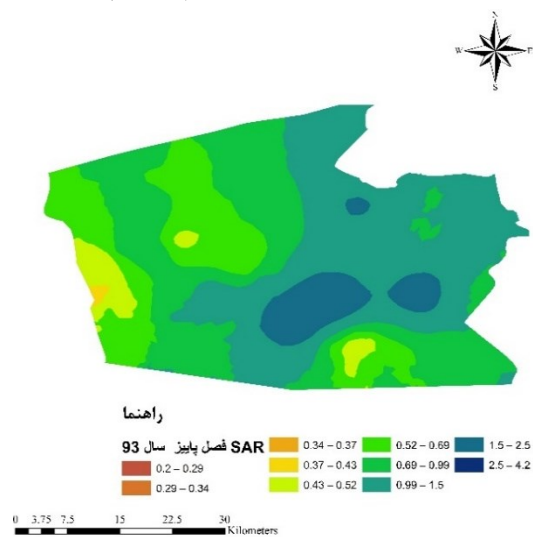
شکل ۴-۷۷. نسبت جذب سدیم فصل بهار ۹۳



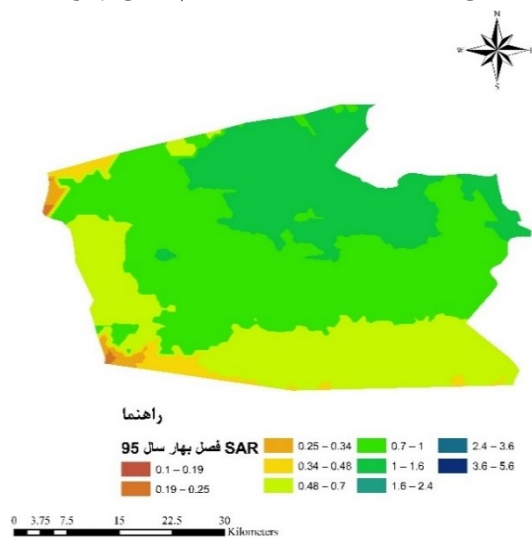
شکل ۴-۷۶. نسبت جذب سدیم فصل پاییز ۹۲



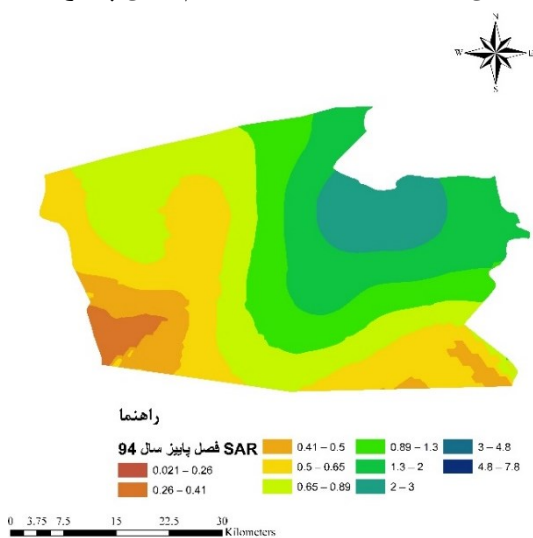
شکل ۴-۷۹. نسبت جذب سدیم فصل بهار ۹۴



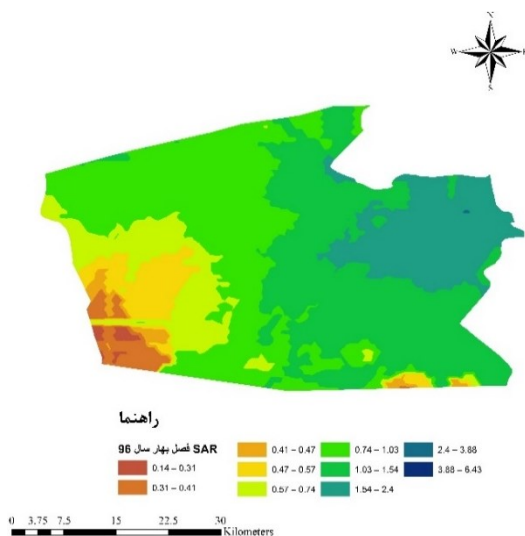
شکل ۴-۷۸. نسبت جذب سدیم فصل پاییز ۹۳



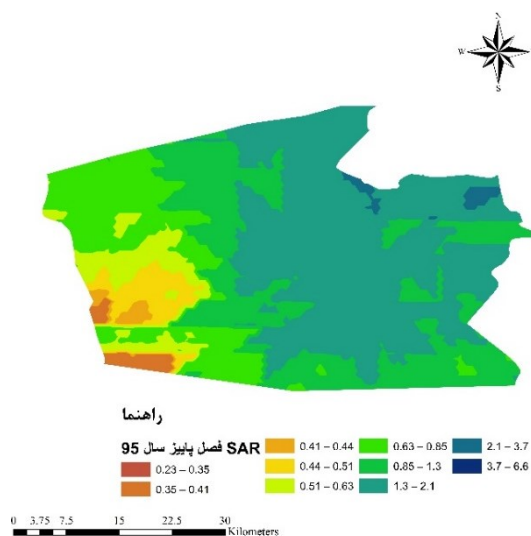
شکل ۴-۸۱. نسبت جذب سدیم فصل بهار ۹۵



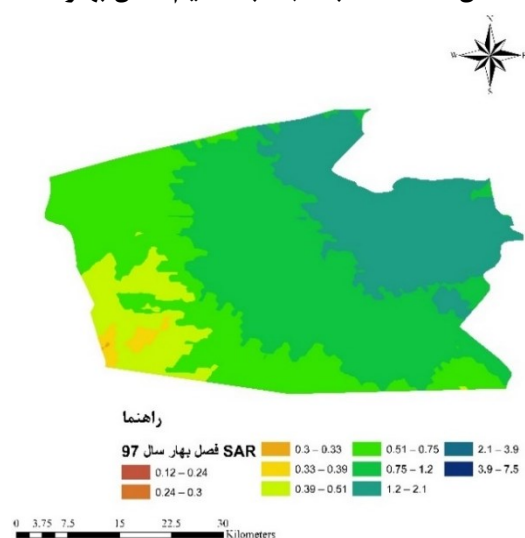
شکل ۴-۸۰. نسبت جذب سدیم فصل پاییز ۹۴



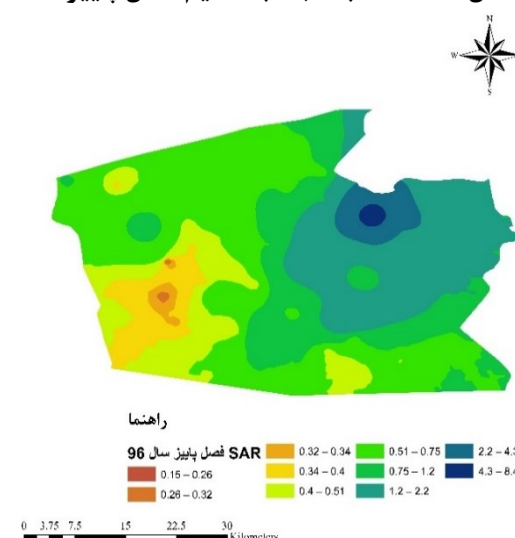
شکل ۴-۸۳. نسبت جذب سدیم فصل بهار ۹۶



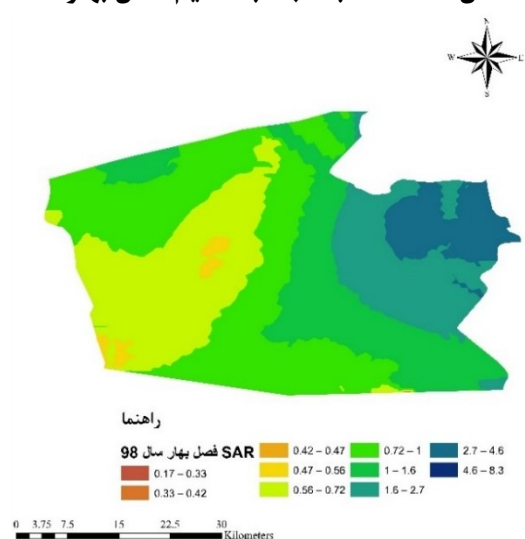
شکل ۴-۸۲. نسبت جذب سدیم فصل پاییز ۹۵



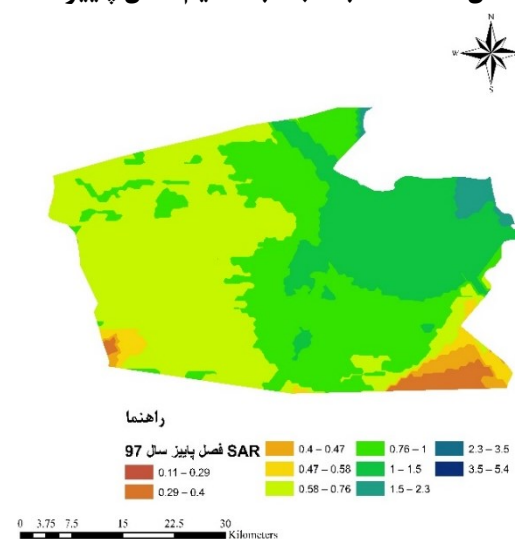
شکل ۴-۸۵. نسبت جذب سدیم فصل بهار ۹۷



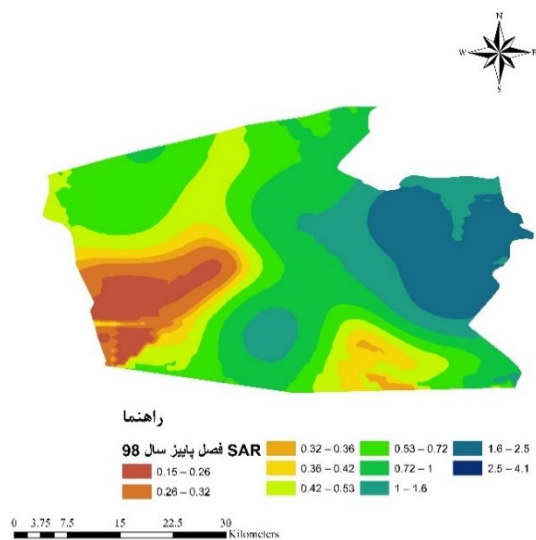
شکل ۴-۸۴. نسبت جذب سدیم فصل پاییز ۹۶



شکل ۴-۸۷. نسبت جذب سدیم فصل بهار ۹۸

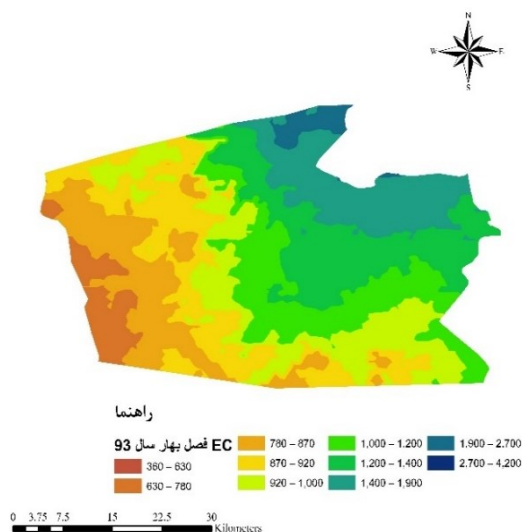


شکل ۴-۸۶. نسبت جذب سدیم فصل پاییز ۹۷

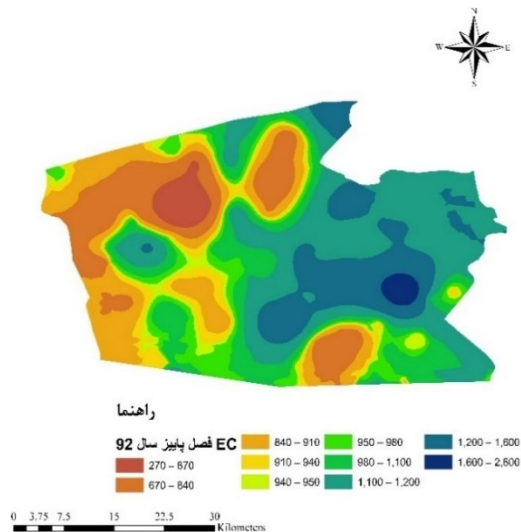


شکل ۴-۸۸. نسبت جذب سدیم فصل پاییز ۹۸

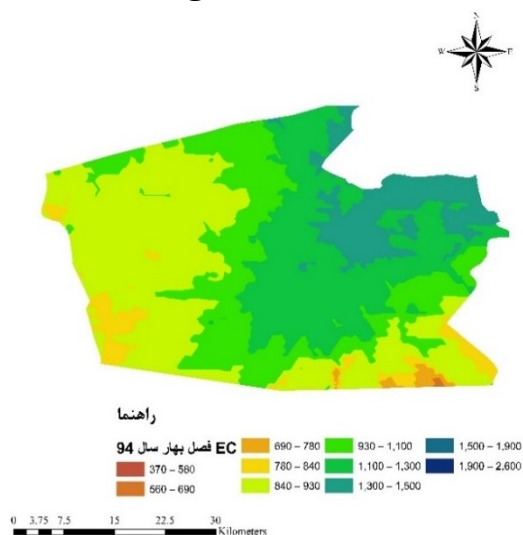
در شکل ۴-۷۶ تا ۴-۸۸ به دلیل الگوی پیچیده، روند خاصی را نمی‌توان در نظر گرفت. لازم به ذکر است در بین این اشکال، فصل پاییز ۹۲ و ۹۳ دارای الگویی مشابه یکدیگر هستند.



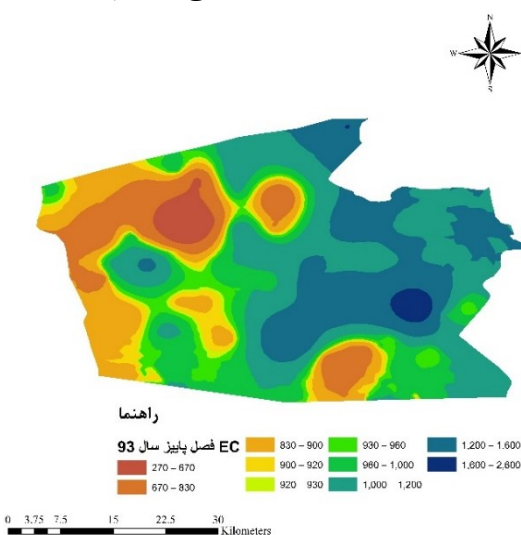
شکل ۴-۹۰. هدایت الکتریکی فصل بهار ۹۳



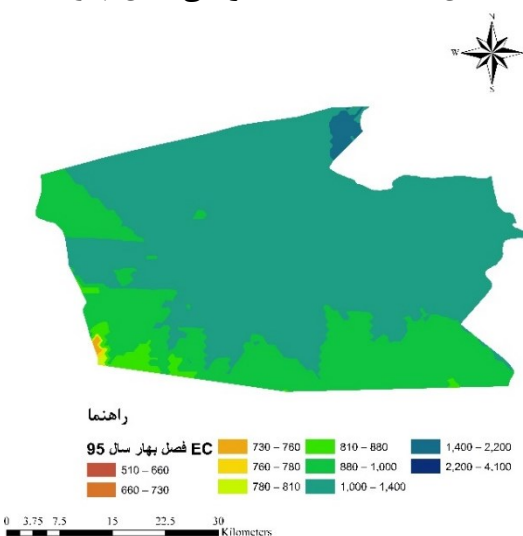
شکل ۴-۸۹. هدایت الکتریکی فصل پاییز ۹۲



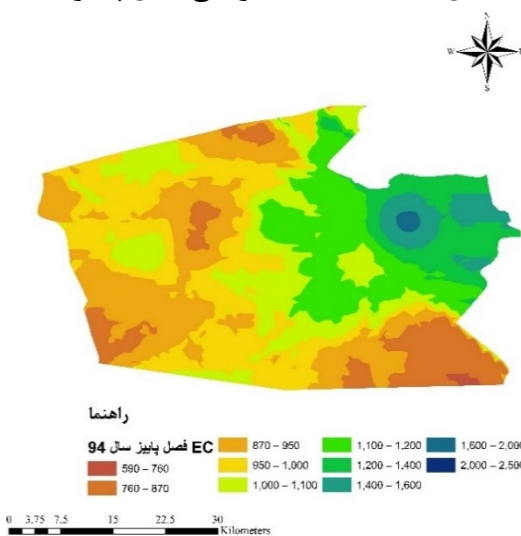
شکل ۴-۹۲. هدایت الکتریکی فصل بهار ۹۴



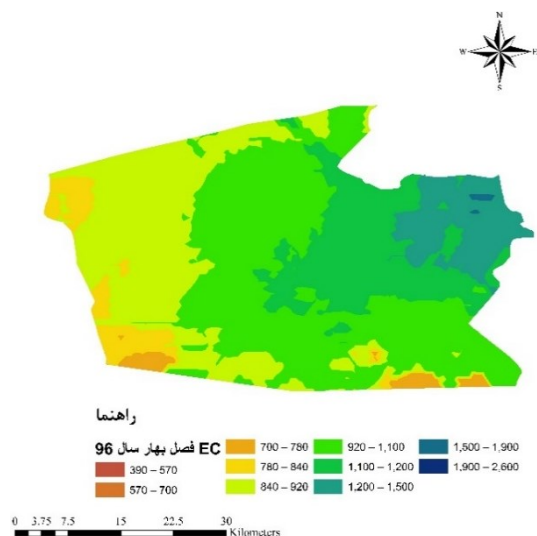
شکل ۴-۹۱. هدایت الکتریکی فصل پاییز ۹۳



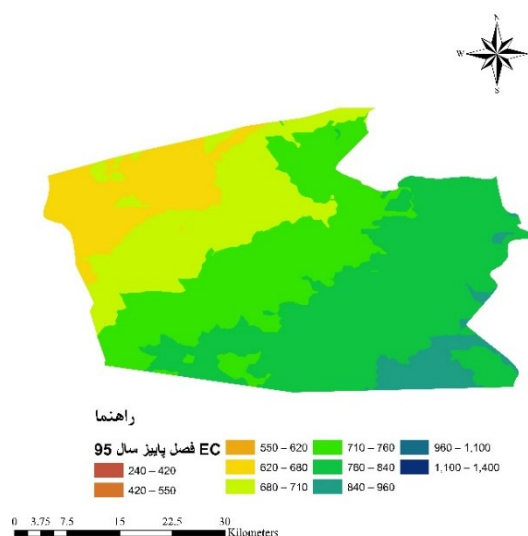
شکل ۴-۹۴. هدایت الکتریکی فصل بهار ۹۵



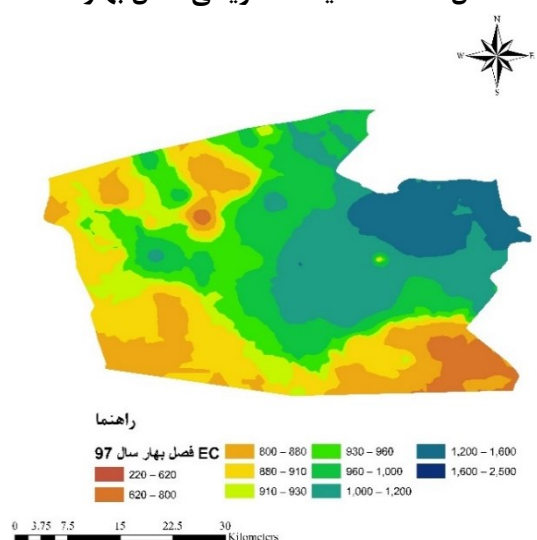
شکل ۴-۹۳. هدایت الکتریکی فصل پاییز ۹۴



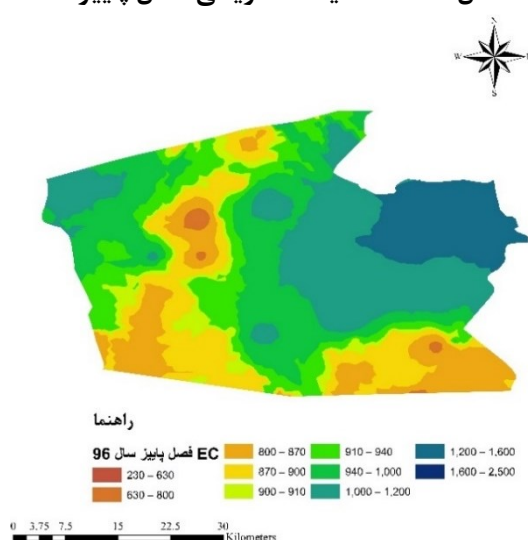
شکل ۴-۹۶. هدایت الکتریکی فصل بهار ۹۶



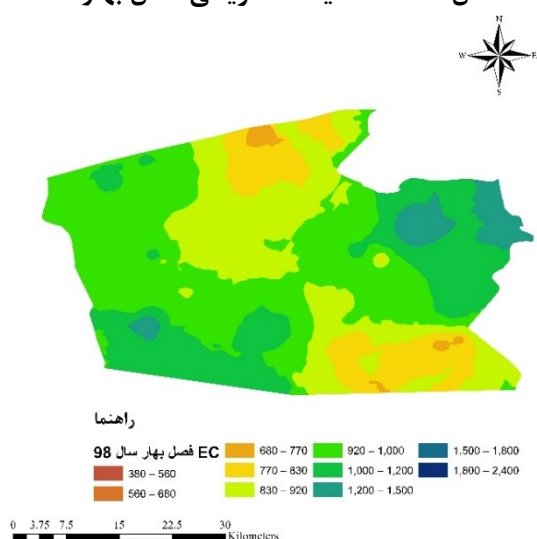
شکل ۴-۹۵. هدایت الکتریکی فصل پاییز ۹۵



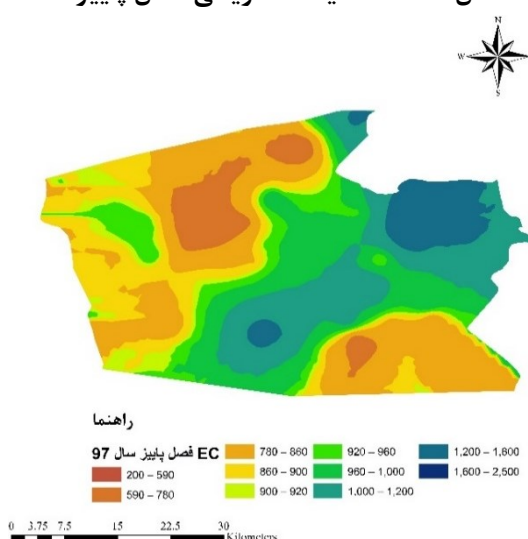
شکل ۴-۹۸. هدایت الکتریکی فصل بهار ۹۷



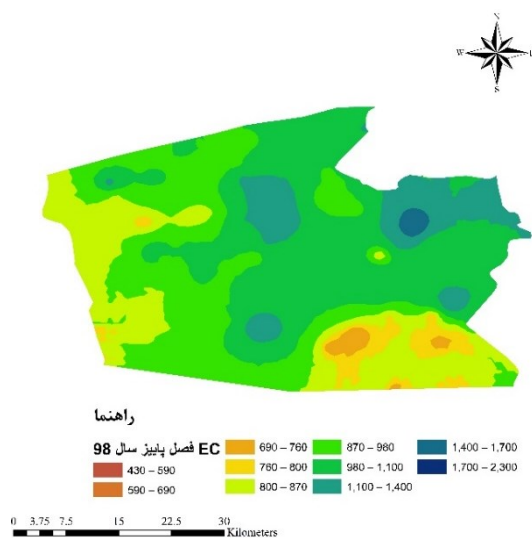
شکل ۴-۹۷. هدایت الکتریکی فصل پاییز ۹۶



شکل ۴-۱۰۰. هدایت الکتریکی فصل بهار ۹۸

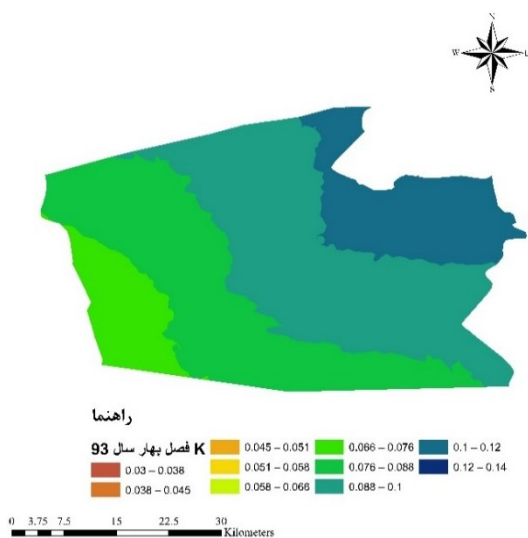


شکل ۴-۹۹. هدایت الکتریکی فصل پاییز ۹۷

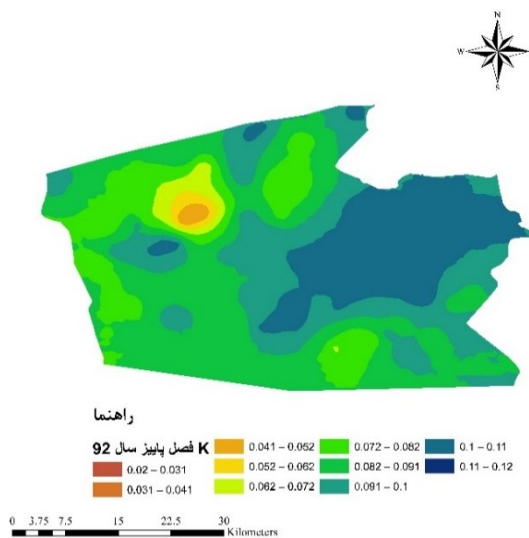


شکل ۴-۱۰۱. هدایت الکتریکی فصل پاییز ۹۸

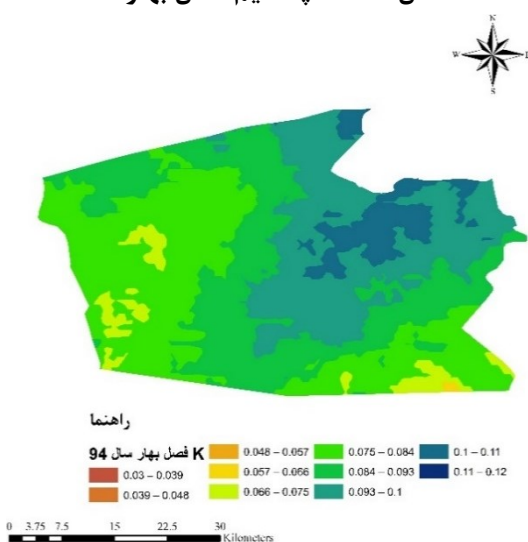
از مشاهده شکل ۴-۸۹ تا ۴-۱۰۱ می‌توان نتیجه گرفت که از مرکز به شرق میزان هدایت الکتریکی افزوده می‌شود. البته در اشکال ۴-۹۴ و ۴-۱۰۰ رفتار آن‌ها با سایر متفاوت می‌باشد.



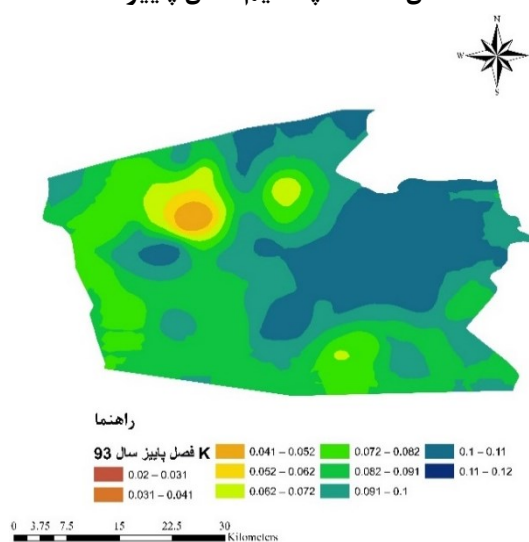
شکل ۴-۱۰۳. پتاسیم فصل بهار ۹۳



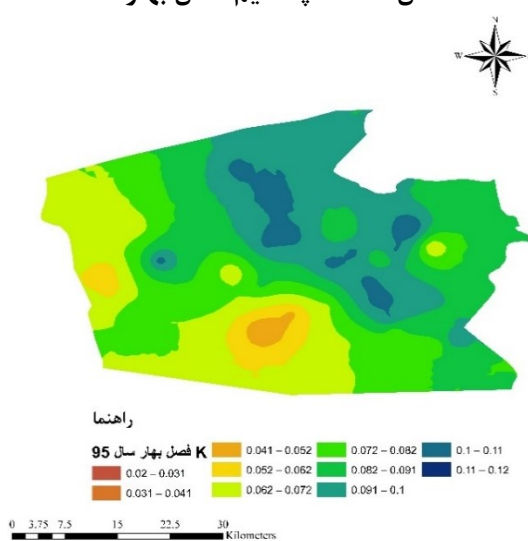
شکل ۴-۱۰۲. پتاسیم فصل پاییز ۹۲



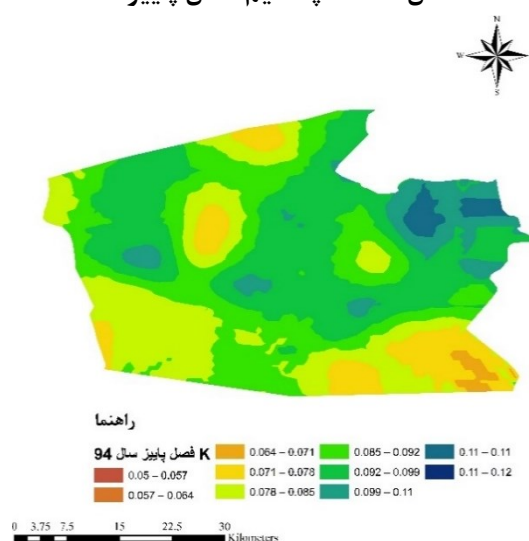
شکل ۴-۱۰۵. پتاسیم فصل بهار ۹۴



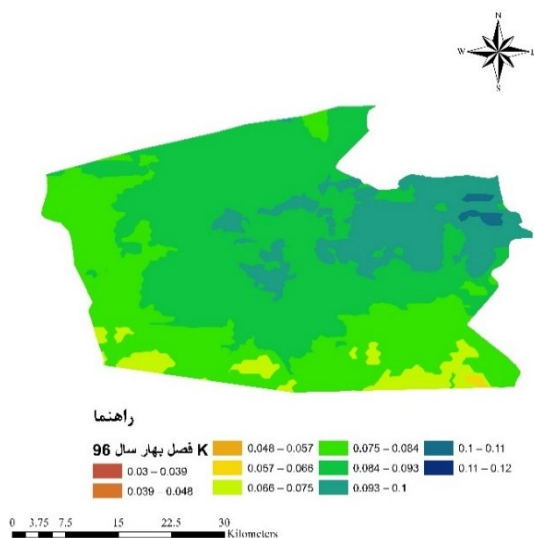
شکل ۴-۱۰۴. پتاسیم فصل پاییز ۹۳



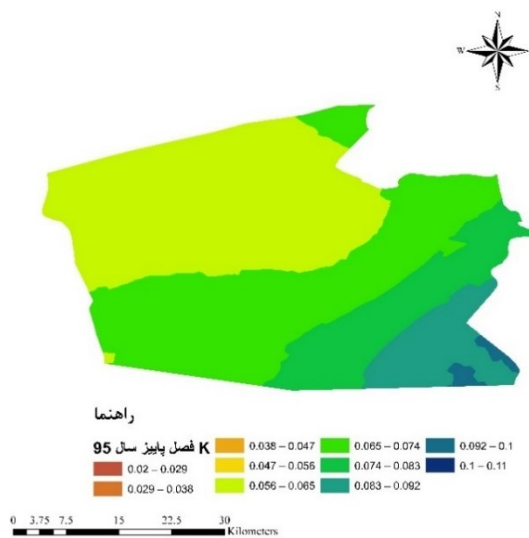
شکل ۴-۱۰۷. پتاسیم فصل بهار ۹۵



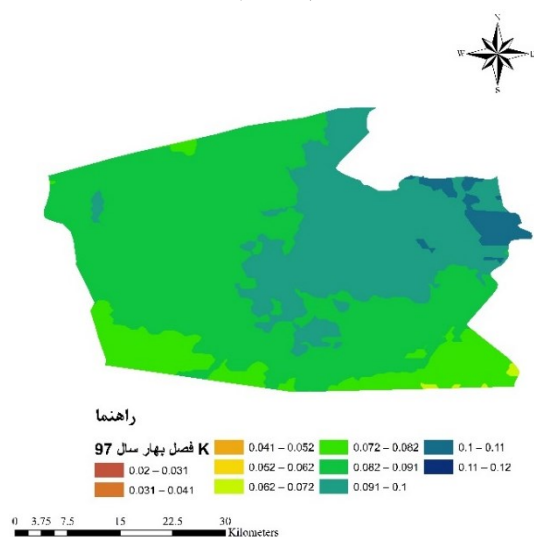
شکل ۴-۱۰۶. پتاسیم فصل پاییز ۹۴



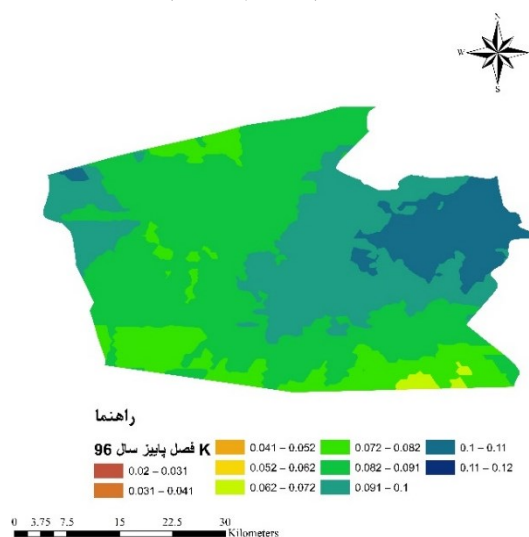
شکل ۴-۱۰۹. پتاسیم فصل بهار ۹۶



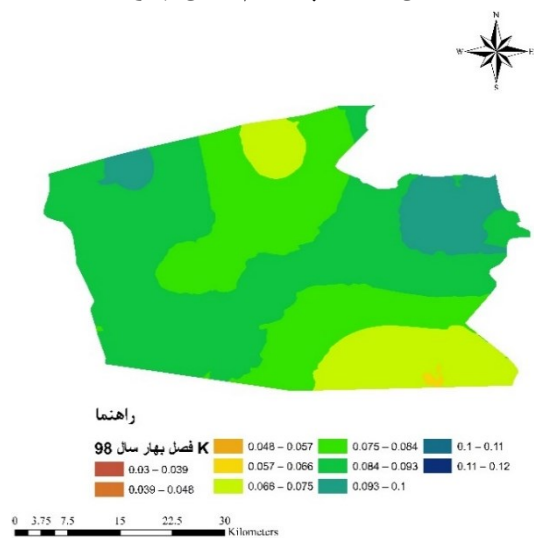
شکل ۴-۱۰۸. پتاسیم فصل پاییز ۹۵



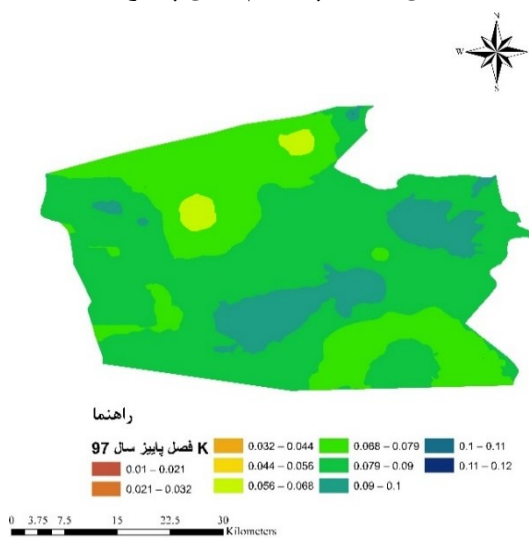
شکل ۴-۱۱۱. پتاسیم فصل بهار ۹۷



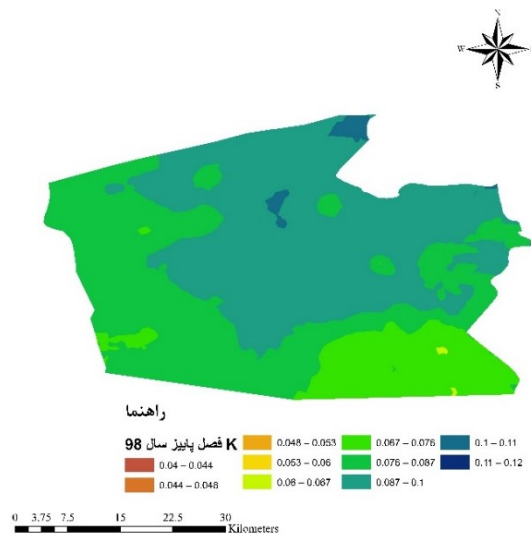
شکل ۴-۱۱۰. پتاسیم فصل پاییز ۹۶



شکل ۴-۱۱۳. پتاسیم فصل بهار ۹۸



شکل ۴-۱۱۲. پتاسیم فصل پاییز ۹۷



شکل ۴-۱۱۴. پتاسیم فصل پاییز ۹۸

باتوجه به داده‌ها به دلیل عدم تبعیت بیشتر داده‌ها با سختی موقت و دائم، می‌توان این نتیجه را گرفت که برای مدلسازی از حالت غیرخطی استفاده نمود. در ادامه به مدلسازی در شبکه پیشرو و ماشین بردار پشتیبان پرداخته شده است.

۴-۴ معیار عملکرد

برای ارزیابی عملکرد مدل تهیه شده و انتخاب مناسب‌ترین مدل از پارامترهای $RMSE$ ، R و R^2 استفاده شد که روابط آن به صورت زیر قابل بیان است.

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (M_i - P_i)^2}{N}} \quad (۱-۴)$$

$$R = \frac{\sum_{i=1}^N (M_i - \bar{M})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^N (M_i - \bar{M})^2} \sqrt{\sum_{i=1}^N (P_i - \bar{P})^2}} \quad (۲-۴)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (M_i - P_i)^2}{\sum_{i=1}^N (P_i)^2} \quad (۳-۴)$$

از آنجا، N تعداد نقاط داده‌ها، M مقدار مشاهده شده، P مقدار پیش بینی شده و $\bar{M}$ و $\bar{P}$ به ترتیب میانگین مقادیر M و P می‌باشند. بهترین تناسب بین سختی موقت و دائم مشاهده شده و پیش بینی شده در آب‌های زیرزمینی زمانی است که RMSE برابر 0 و R^2 برابر 1 باشد [۴۳].

۴-۵ مدل‌سازی به روش شبکه عصبی پیشرو

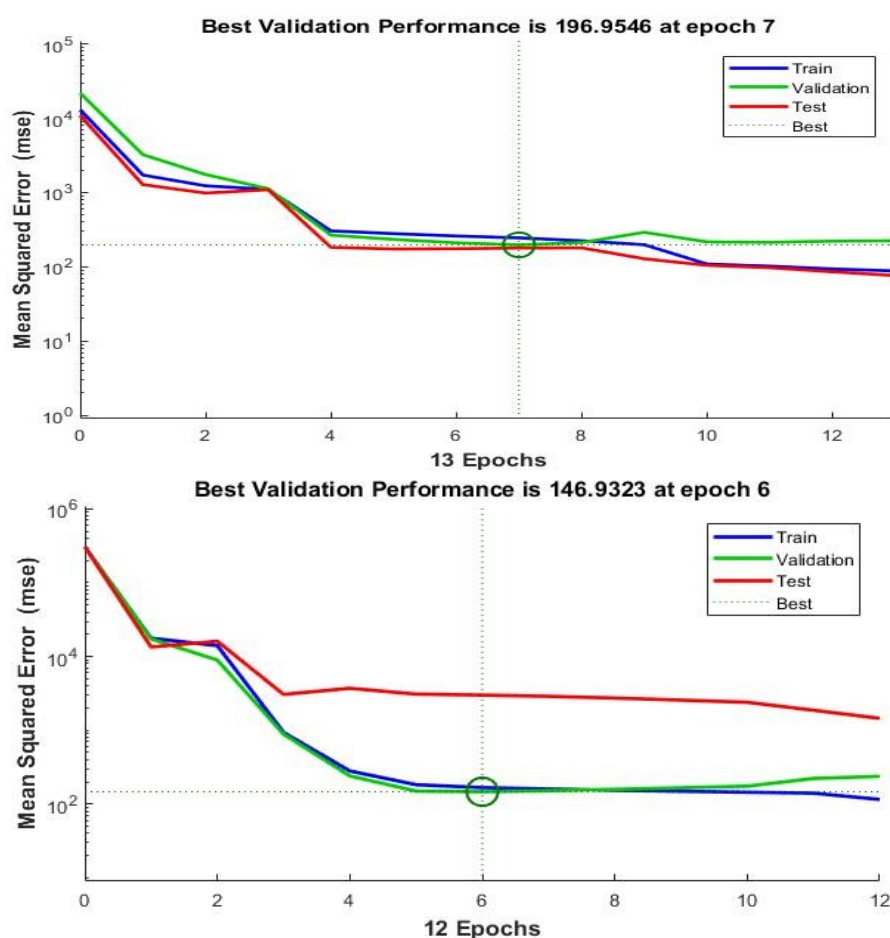
هدف از این مطالعه، ساخت مدلی بهینه، به طوری که توانایی پیش بینی سختی موقت و دائم آب به کمک داده‌هایی که به راحتی قابل اندازه گیری هستند، می‌باشد. برای این منظور براساس داده‌های موجود مدلی را آموزش داده که توانایی پیش بینی را با بهترین تقریب ممکن داشته باشد. به منظور مدل‌سازی شبکه عصبی پیشرو از نرم افزار MATLAB استفاده شده است. برای مدل‌سازی در نرم افزار، ابتدا داده‌های ورودی و خروجی را مشخص، سپس آن‌ها را نرمال‌سازی کرده و با استفاده از تابع آموزش و تابع یادگیری مناسب مدل ساخته و اجرا می‌گردد. بنابراین در صورتی که بتوان براساس داده‌های موجود، تابع مناسبی را جهت دریافت ورودی‌های عنوان شده ایجاد کرد، آموزش مدل به درستی انجام شده و از آن می‌توان برای پیش‌بینی سختی دائم و موقت در سایر نقاط استفاده نمود.

۴-۶ تعیین الگوریتم آموزش و یادگیری

باتوجه به استفاده از شبکه عصبی پیشرو در این مطالعه، لازم است تابع یادگیری و آموزش مورد بررسی قرار گیرند. در تحقیقات زیادی کارایی انواع تابع یادگیری و آموزش مورد مطالعه قرار گرفته است. از میان توابع موجود برای یادگیری شبکه، گرادیان کاهشی با وزن مومنتوم عملکرد مناسبی را نسبت به سایر توابع یادگیری دارد [۱۰]. به علاوه تابع آموزش لونیگ-مارکوارت با ترکیبی از دو الگوریتم روش گرادیان کاهشی و روش گaus-نیوتون را برای شبکه عصبی حل می‌کند که به عنوان الگوریتمی کارآمد در نظر گرفته می‌شود [۱۱].

۴-۷ آموزش و تحلیل شبکه

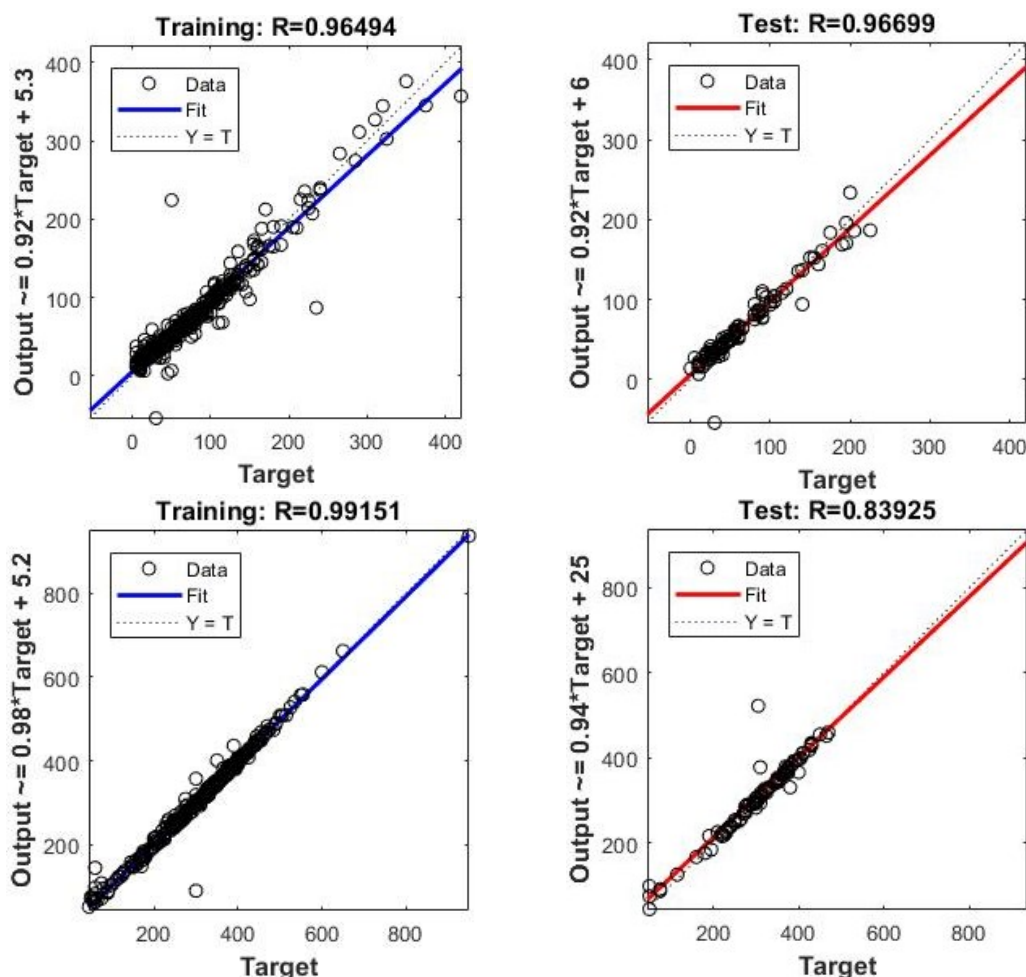
از آنجا که ورودی‌ها و خروجی‌ها تعیین شده‌اند، داده‌ها همانند روش SVM تحت آموزش قرار می‌گیرند. در حین آموزش خطاها از طریق نوروں‌ها با تنظیم وزن آن به عقب منتقل می‌شوند. شبکه عصبی پیشرو دارای ۸ نوروں در لایه اول می‌باشد. شکل ۴-۱۱۵، نمودار فرایند آموزش برای سختی دائم و موقت را در سه مرحله آزمون، اعتبارسنجی و آموزش نشان می‌دهد. مطابق شکل، بهترین خط نشان دهنده بهترین میزان MSE برای شبکه طراحی شده است. فرایند آموزش شبکه هنگامی که مقادیر حاصل از اعتبارسنجی و آزمون نزدیک به هم باشد، قابل قبول می‌باشد.



شکل ۴-۱۱۵. (الف) نمودار عملکرد فرایند آموزش سختی دائم برای رسیدن به کمترین خطا
(ب) نمودار عملکرد فرایند آموزش سختی موقت برای رسیدن به کمترین خطا

همانند مدل SVM، در طول آموزش به صورت تصادفی ۸۰٪ از داده‌های ارائه شده برای آموزش و ۲۰٪ باقیمانده مجموعه داده‌ها برای آزمون و اعتبارسنجی شبکه استفاده شد.

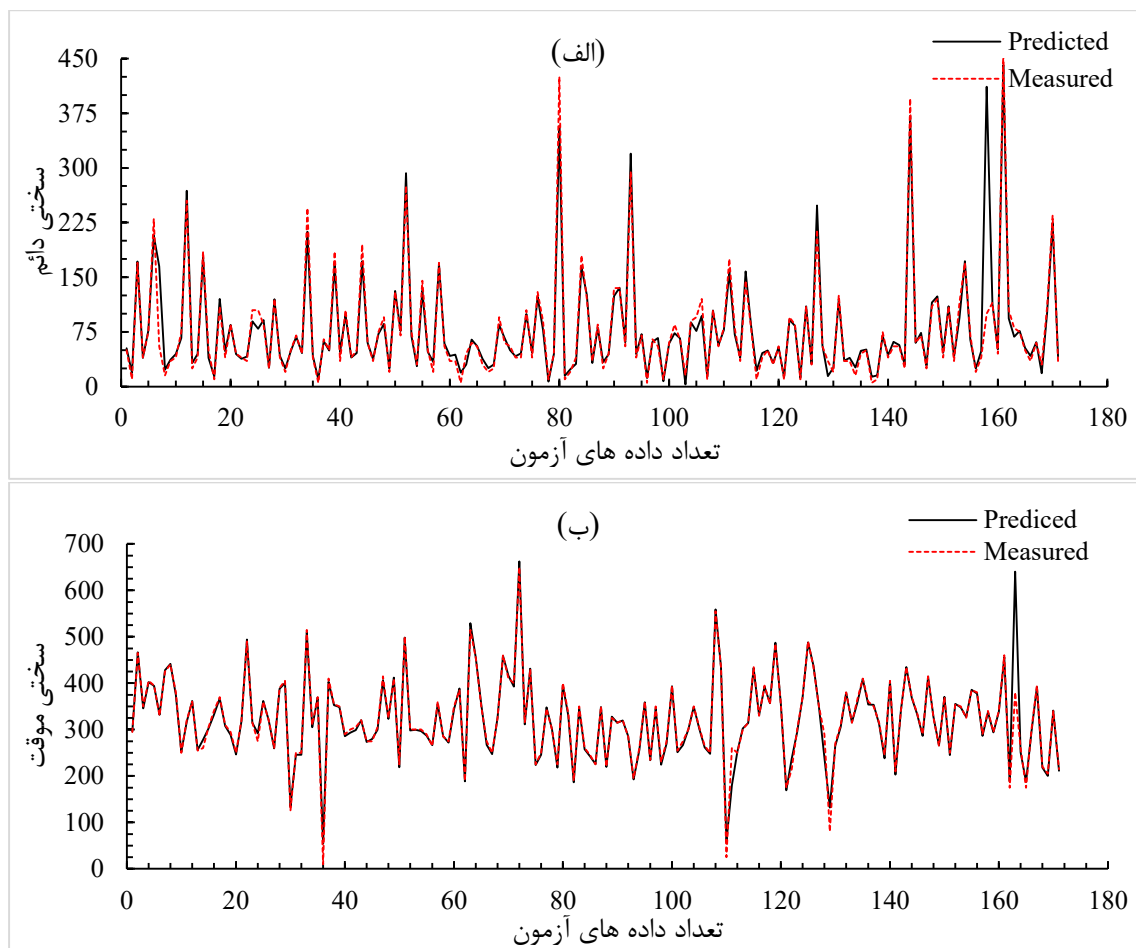
در شکل (الف) ۴-۱۱۵، تحلیل شبکه در مقدار خطای میانگین مربعات ۱۹۶/۹۵ و در تعداد خطای اعتبار سنجی ۷ برای سختی دائم و در شکل (ب) ۴-۱۱۵ مقدار خطای میانگین مربعات ۱۴۶/۹۳ در تعداد خطای اعتبارسنجی ۶ متوقف شد. علت این امر به دلیل جلوگیری از ایجاد بیش برآزش می‌باشد. در حقیقت، در صورت ادامه تحلیل، شبکه داده‌ها را حفظ کرده و قادر به پیش بینی درستی از داده‌ها نخواهد بود.



شکل ۴-۱۱۶. (الف) نمودار رگرسیون بین داده‌های محاسباتی و مشاهداتی به ازای کل داده‌های سختی دائم

(ب) نمودار رگرسیون بین داده‌های محاسباتی و مشاهداتی به ازای کل داده‌های سختی موقت

به منظور بررسی چگونگی تناسب داده‌های ورودی و خروجی از نمودار رگرسیون استفاده شده است. در شکل ۴-۱۱۶ (الف) و ۴-۱۱۶ (ب)، به ترتیب نمودار رگرسیون فرایند آموزش برای سختی دائم و موقت است. در صورتی که نتایج بر روی خط چین قرار داشته باشد نشان دهنده بهترین نتیجه می‌باشد ($R=1$). مطابق شکل، محور افقی نشان دهنده تابع هدف و محور عمودی نشان دهنده خروجی شبکه برای داده‌های آموزش و آزمون می‌باشد. مطابق شکل فوق، نتایج بدست آمده برای داده‌های آزمون و آموزش به ترتیب ۰/۹۶۶۹ و ۰/۹۶۴۹ برای سختی دائم و ۰/۸۳۹۲ و ۰/۹۹۱۵ برای سختی موقت می‌باشد.



شکل ۴-۱۱۷. (الف) مقایسه غلظت سختی دائم پیش بینی شده و اندازه‌گیری در مرحله آزمون مدل ANN
(ب) مقایسه غلظت سختی موقت پیش بینی شده و اندازه‌گیری در مرحله آزمون مدل ANN

شکل ۴-۱۱۷ مقایسه مقادیر غلظت سختی دائم و موقت مشاهداتی و پیش بینی شده برای ۱۷۱ نمونه تحت آزمون برای مدل ANN را نشان می‌دهد. در قسمت‌هایی از نمودار مقداری خطا وجود دارد، با این حال ارزیابی نشان از قرار داشتن خطوط بر روی یکدیگر را دارد. در نهایت باتوجه به مدلسازی انجام شده، نتایج برای سختی دائم و موقت در جدول ۴-۲ گزارش شده است.

جدول ۴-۳. نتایج شبکه عصبی Feedforward سختی دائم

Feedforward Back-Propagation		
نتایج مدل	آموزش	آزمون
R (Correlation_coefficient)	۰/۹۶۴۹	۰/۹۶۶۹
R^2	۰/۹۷۰۹	۰/۹۳۰۳
MSE	۲۲۷/۵۵	۷۴۸/۷۶
RMSE(mg/L)	۱۵/۰۸۴	۲۷/۳۶

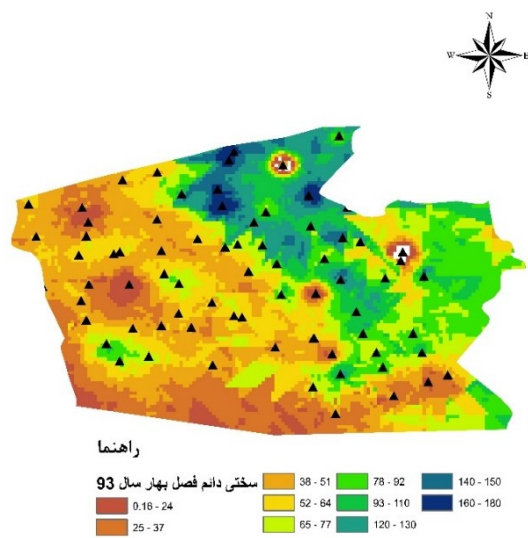
جدول ۴-۴. نتایج شبکه عصبی Feedforward سختی موقت

Feedforward Back-Propagation		
نتایج مدل	آموزش	آزمون
R (Correlation_coefficient)	۰/۹۹۱۵	۰/۸۳۹۲
R^2	۰/۹۶۸۹	۰/۹۶۸۸
MSE	۵۹۰/۸۷	۵۰۴/۷۸
RMSE(mg/L)	۲۴/۳۰	۲۲/۴۶

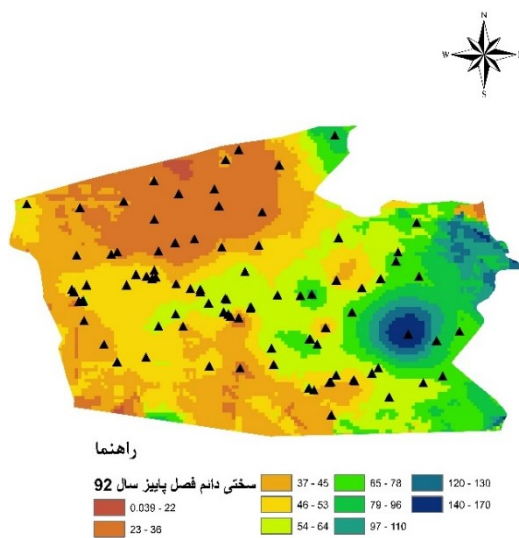
۴-۸ پیش بینی سختی موقت و دائم آب به کمک شبکه عصبی

پیشرو

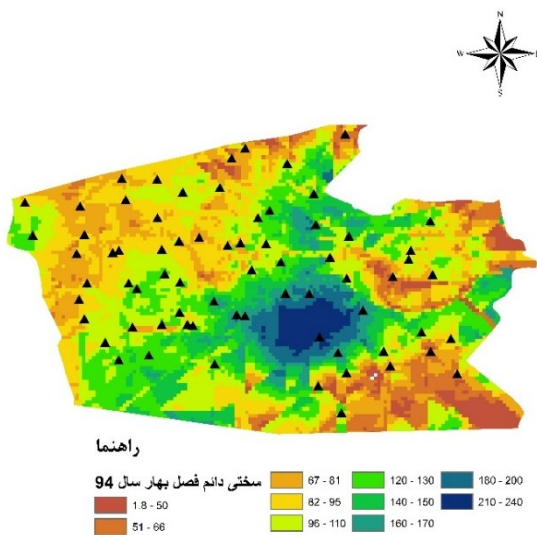
پس از ایجاد مدل بهینه، با استفاده از داده‌های بدست آمده از درون‌یابی در ArcGIS، مقادیر را به صورت یک ماتریس ۸×۴۷۷۵۶۸ به مدل داده و اقدام به پیش بینی می‌کنیم. در نهایت خروجی آن که یک ماتریس با سطر ثابت و ستون یک می‌باشد را به ArcGIS داده و نتایج را به صورت گرافیکی نمایش می‌دهیم.



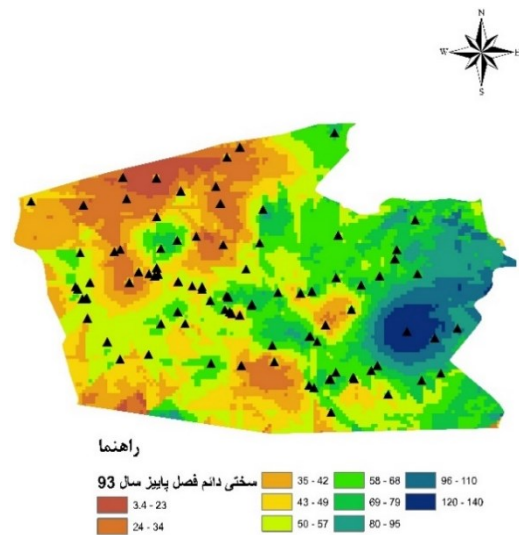
شکل ۴-۱۱۹. سختی دائم بهار ۹۳



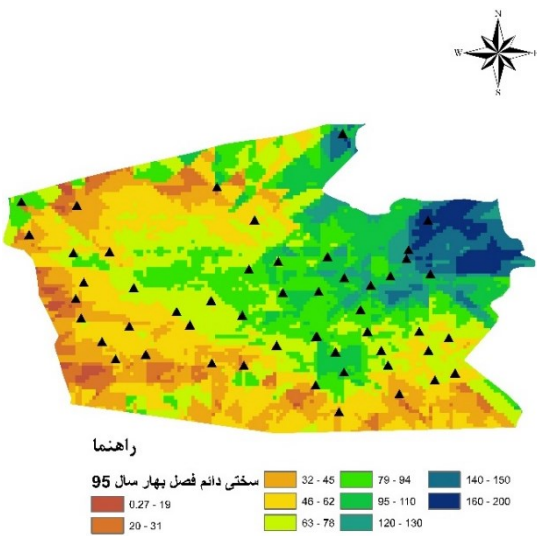
شکل ۴-۱۱۸. سختی دائم پاییز ۹۲



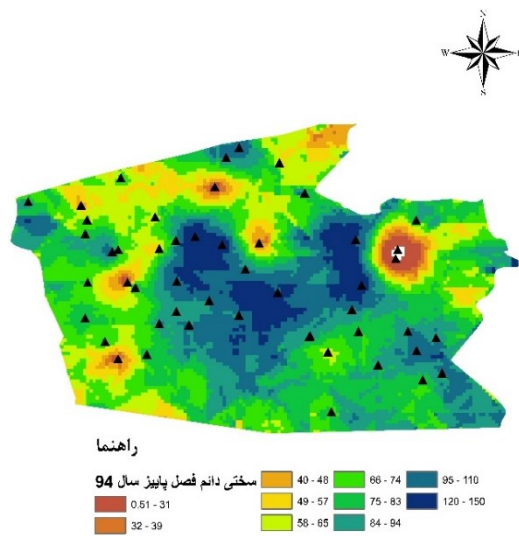
شکل ۴-۱۲۱. سختی دائم بهار ۹۴



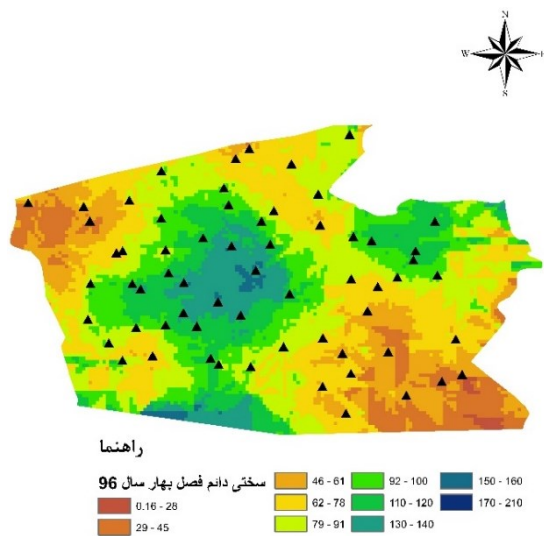
شکل ۴-۱۲۰. سختی دائم پاییز ۹۳



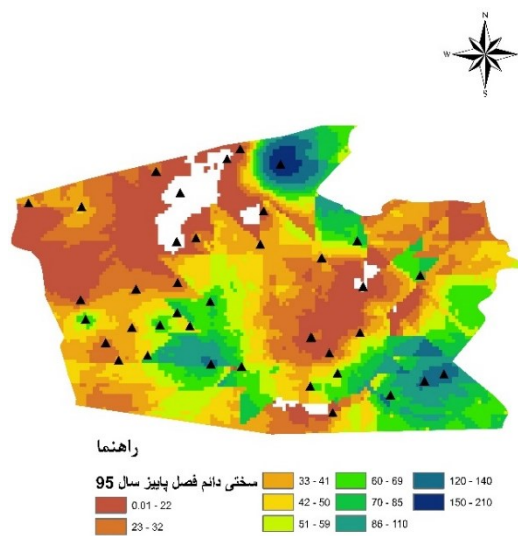
شکل ۴-۱۲۳. سختی دائم بهار ۹۵



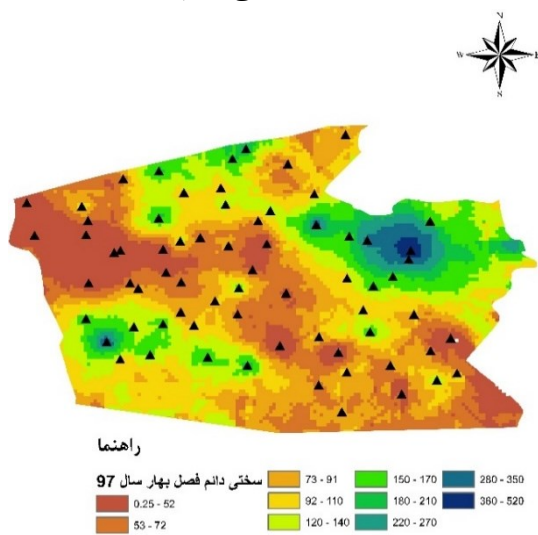
شکل ۴-۱۲۲. سختی دائم پاییز ۹۴



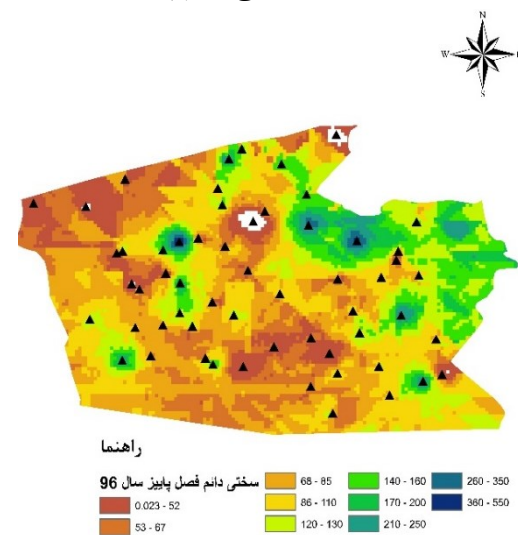
شکل ۴-۱۲۵. سختی دائم بهار ۹۶



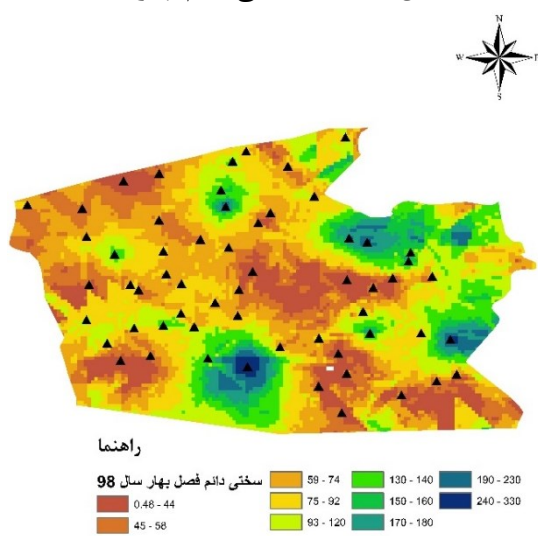
شکل ۴-۱۲۴. سختی دائم پاییز ۹۵



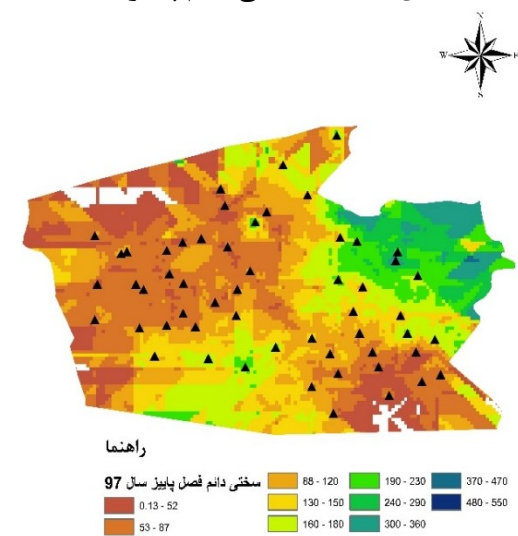
شکل ۴-۱۲۷. سختی دائم بهار ۹۷



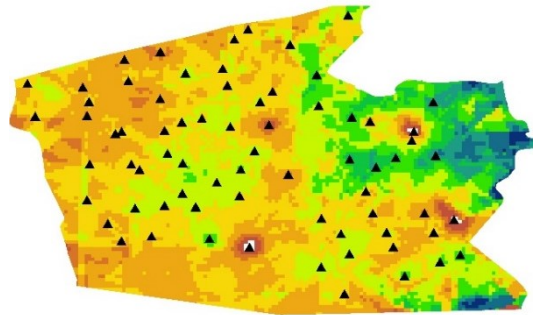
شکل ۴-۱۲۶. سختی دائم پاییز ۹۶



شکل ۴-۱۲۹. سختی دائم بهار ۹۸



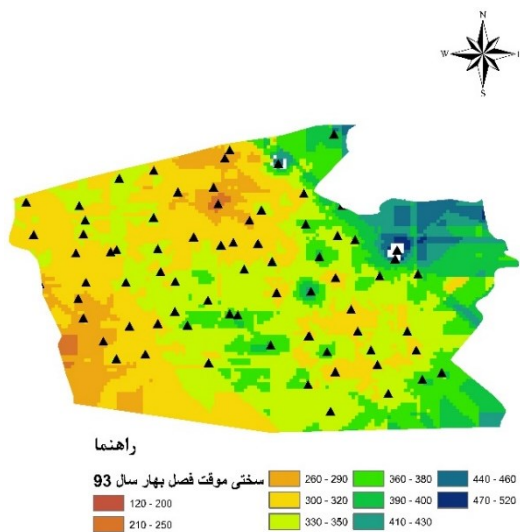
شکل ۴-۱۲۸. سختی دائم پاییز ۹۷



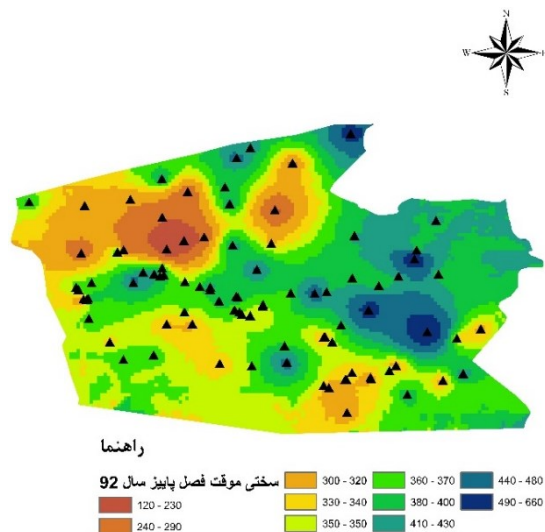
راهنما

سختی دائم فصل پاییز سال ۹۸	60 - 83	140 - 160	210 - 240
0.43 - 38	84 - 110	170 - 180	250 - 300
39 - 59	120 - 130	190 - 200	

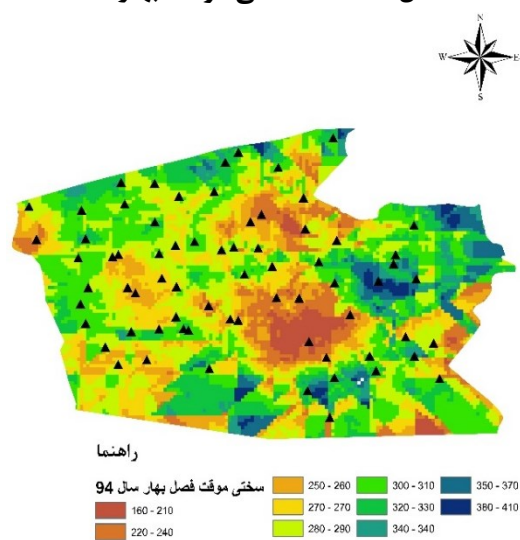
شکل ۴-۱۳۰. سختی دائم پاییز ۹۸



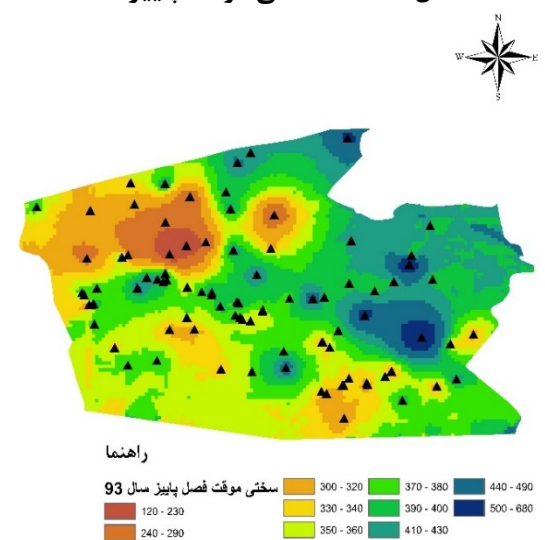
شکل ۴-۱۳۲. سختی موقت فصل بهار ۹۳



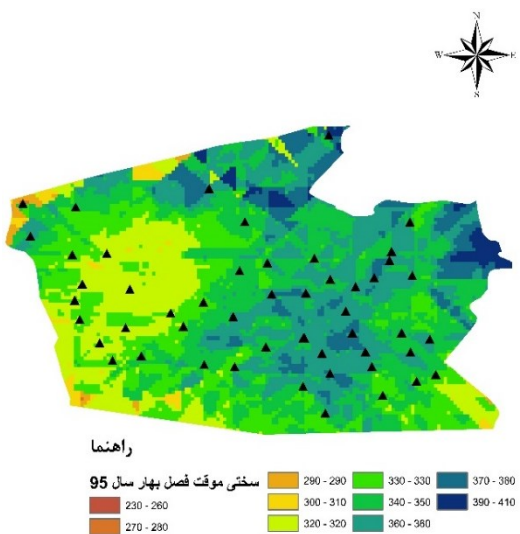
شکل ۴-۱۳۱. سختی موقت فصل پاییز ۹۲



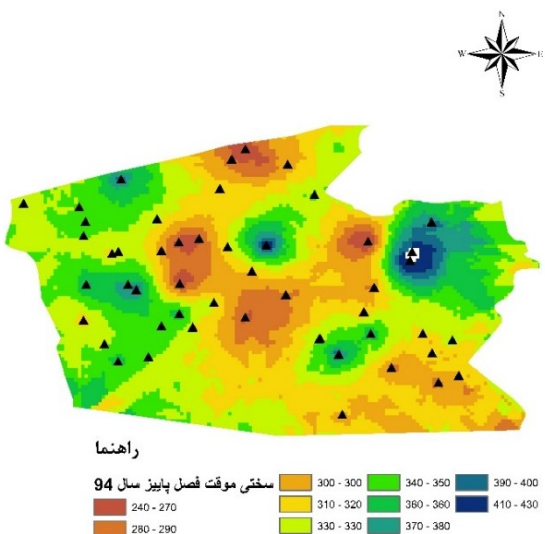
شکل ۴-۱۳۴. سختی موقت فصل بهار ۹۴



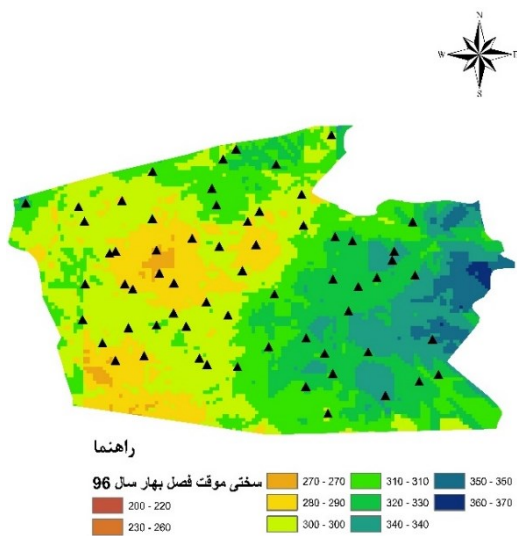
شکل ۴-۱۳۳. سختی موقت فصل پاییز ۹۳



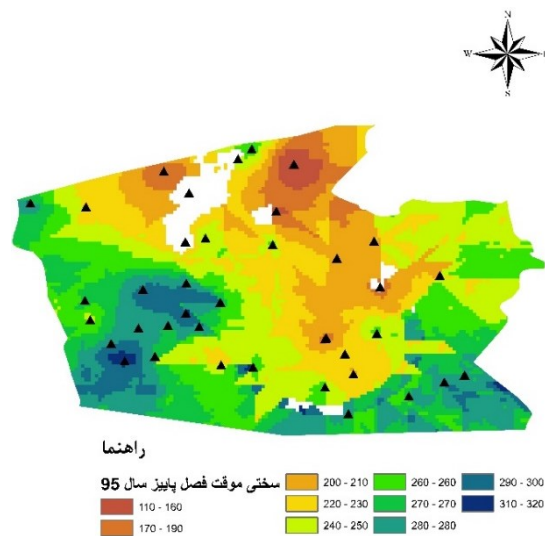
شکل ۴-۱۳۶. سختی موقت فصل بهار ۹۵



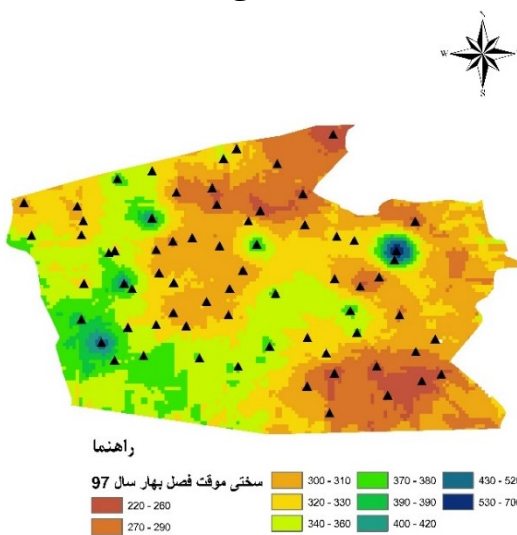
شکل ۴-۱۳۵. سختی موقت فصل پاییز ۹۴



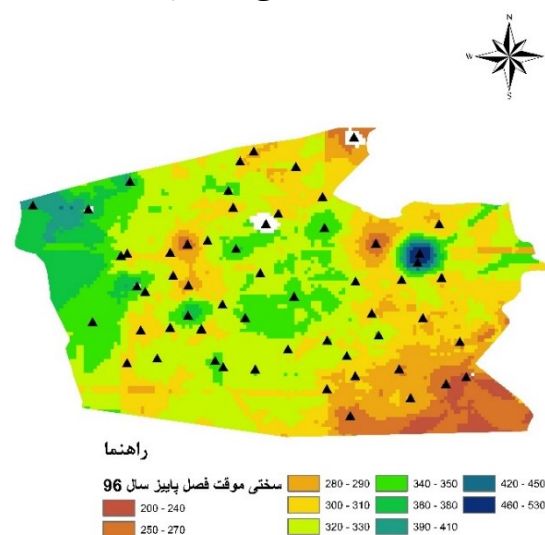
شکل ۴-۱۳۸. سختی موقت بهار ۹۶



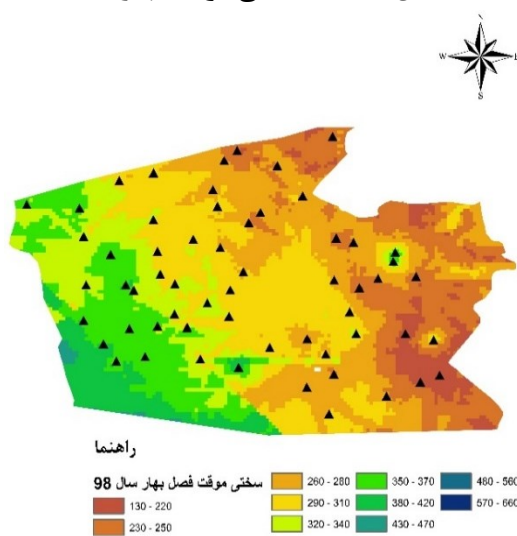
شکل ۴-۱۳۷. سختی موقت پاییز ۹۵



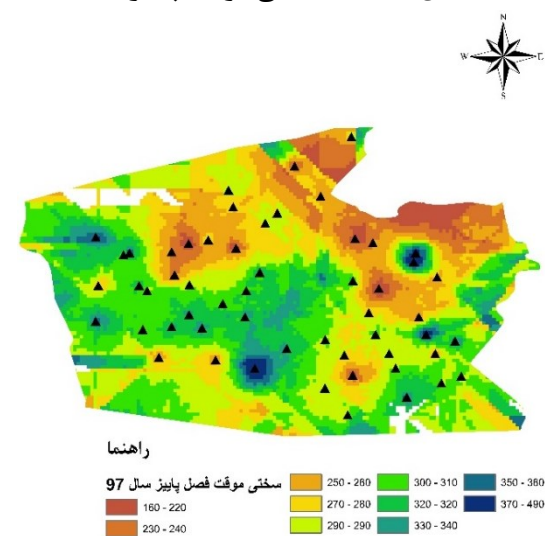
شکل ۴-۱۴۰. سختی موقت بهار ۹۷



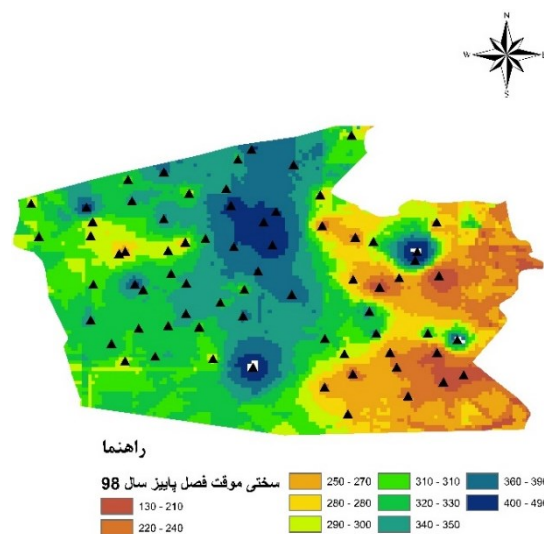
شکل ۴-۱۳۹. سختی موقت پاییز ۹۶



شکل ۴-۱۴۲. سختی موقت بهار ۹۸



شکل ۴-۱۴۱. سختی موقت پاییز ۹۷



شکل ۴-۱۴۳. سختی موقت پاییز ۹۸

لازم ذکر است در قسمت‌هایی از نقشه، شبکه عصبی قادر به محاسبه مقادیر سختی دائم و موقت نمی‌باشد و بصورت پیکسل‌های خالی نشان داده شد.

۹-۴ مدل‌سازی به روش رگرسیون بردار پشتیبان

هدف از مدل‌سازی به روش بردار پشتیبان، ساخت مدلی مناسب که توانایی پیش‌بینی غلظت سختی موقت و دائم توسط داده‌های قابل اندازه‌گیری را داشته باشد. بر این اساس، با استفاده از داده‌های موجود، مدلی آموزش داده می‌شود که توانایی پیش‌بینی با دقت مناسب را داشته باشد. در این مطالعه از تابع `fitrsvm` که توسط متلب نوشته شده، برای تهیه مدل استفاده گردید. بنابراین برای مدل‌سازی، ابتدا داده‌های ورودی و خروجی را مشخص، سپس آن‌ها را نرمال‌سازی می‌کنیم. در نهایت با انتخاب پارامترهای مناسب پارامترهای بهینه و تابع کرنل، مدل اجرا می‌گردد.

۱۰-۴ نرمالسازی داده‌ها

ماشین بردار پشتیبان یک الگوریتم یادگیری ماشین است که بر پایه کمینه سازی ریسک ساختاری برای تعمیم بهتر تعداد نمونه‌ها می‌باشد. ماشین بردار پشتیبان توانایی حل مسائل طبقه بندی و رگرسیون را دارد که برای مسائل رگرسیون به آن رگرسیون بردار پشتیبان می‌نامند. هدف SVR، تخمین یک تابع با تقریب بالا برای مدل پیش بینی می‌باشد. [۴۴] برای مدل سازی ابتدا باید پیش پردازش صورت گیرد. یکی از روش‌های پیش پردازش، نرمالسازی می‌باشد. برای جلوگیری از وابستگی به واحد اندازه‌گیری، داده‌ها باید نرمالسازی شوند. این شامل تبدیل داده‌ها به محدوده کوچکتر یا مشترک مانند (۱،-۱) یا (۱،۰) می‌باشد. نرمالسازی داده‌ها تلاش می‌کند تا به همه ویژگی‌ها وزنی برابر بدهد. روش‌های زیادی برای نرمالسازی داده‌ها وجود دارد که در اینجا از نرمالسازی Min-Max استفاده شده است که رابطه آن به صورت زیر بیان شد.

$$X_{normalized} = \frac{x - x_{minimum}}{x_{maximum} - x_{minimum}} \quad (4-4)$$

این تبدیلات خطی بوده و فقط مقیاس داده تغییر می‌کند. در این تحقیق با استفاده از کد نویسی در MATLAB تمامی داده‌های ورودی و خروجی به بازه (۱،۰) نرمال شده است [۴۵].

۱۱-۴ تعیین تابع کرنل

در این مطالعه باتوجه به استفاده از رگرسیون غیرخطی، لازم است تابع کرنل مناسبی مطابق توابع جدول ۳-۱ انتخاب گردد. در تحقیقات زیادی کارایی انواع تابع مورد بررسی قرار گرفت که بیشتر آن‌ها کارایی تابع RBF را نسبت به توابع دیگر نشان دادند [۴۴].

$$e^{-\sigma \|x - x'\|^2} \quad (4-5)$$

۴-۱۲ مدل‌سازی سختی موقت و دائم به کمک SVM

باتوجه به موارد ذکر شده، بهینه ترین مدل برای تخمین سختی دائم و موقت در دشت بابل-آمل به کمک داده های موجود در جدول ۴-۴ و ۴-۵ پارامترهای آن گزارش شده است. مدل مورد نظر توسط ۶۸۴ داده آموزش دیده شده و با ۱۷۱ داده مورد ارزیابی قرار گرفته است. مطابق جدول، اختلاف کم در مرحله آموزش و آزمون نشان از ارزیابی خوب داده‌های موجود در این مطالعه را دارد.

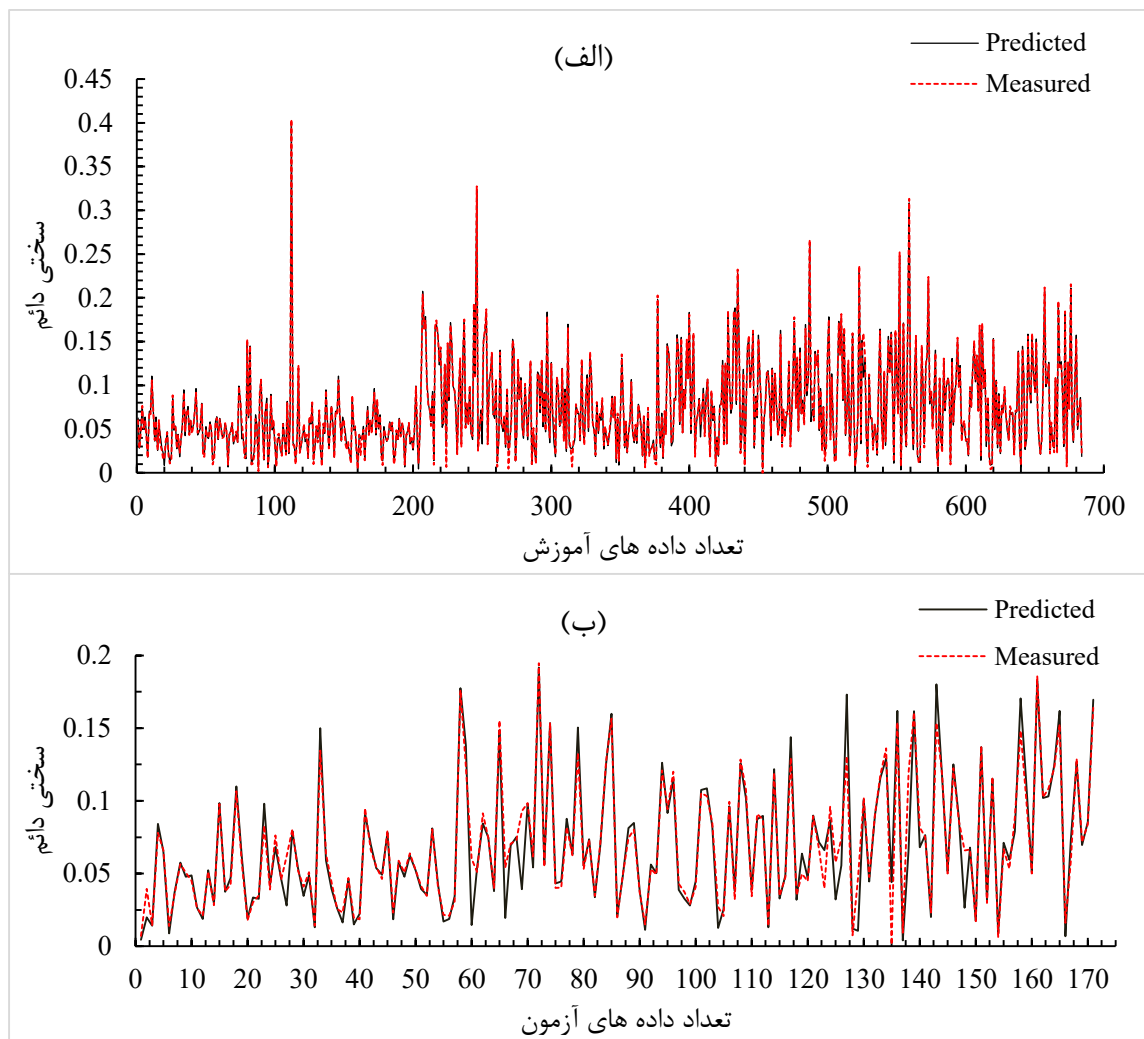
جدول ۴-۵. نتایج مدل Support Vector سختی دائم

پارامتر بهینه	C	$gamma$
	۰/۵	۲/۵
$\mathcal{E}$	۰/۰۰۴۷	
نتایج مدل	آموزش	آزمون
R^2	۰/۹۹۱	۰/۹۳۰۴
MSE	۰/۰۰۰۰۲۳۵	۰/۰۰۰۱۴۰۱
RMSE(Normalized)	۰/۰۰۴۹	۰/۰۱۱۸

جدول ۴-۶. نتایج مدل Support Vector سختی موقت

پارامتر بهینه	C	$gamma$
	۰/۵	۲/۵
$\mathcal{E}$	۰/۰۰۶۸	
نتایج مدل	آموزش	آزمون
R^2	۰/۹۹۴۴	۰/۹۸۳۳
MSE	۰/۰۰۰۰۸۸۳	۰/۰۰۰۰۳۳۸
RMSE(Normalized)	۰/۰۰۹۴	۰/۰۰۵۸

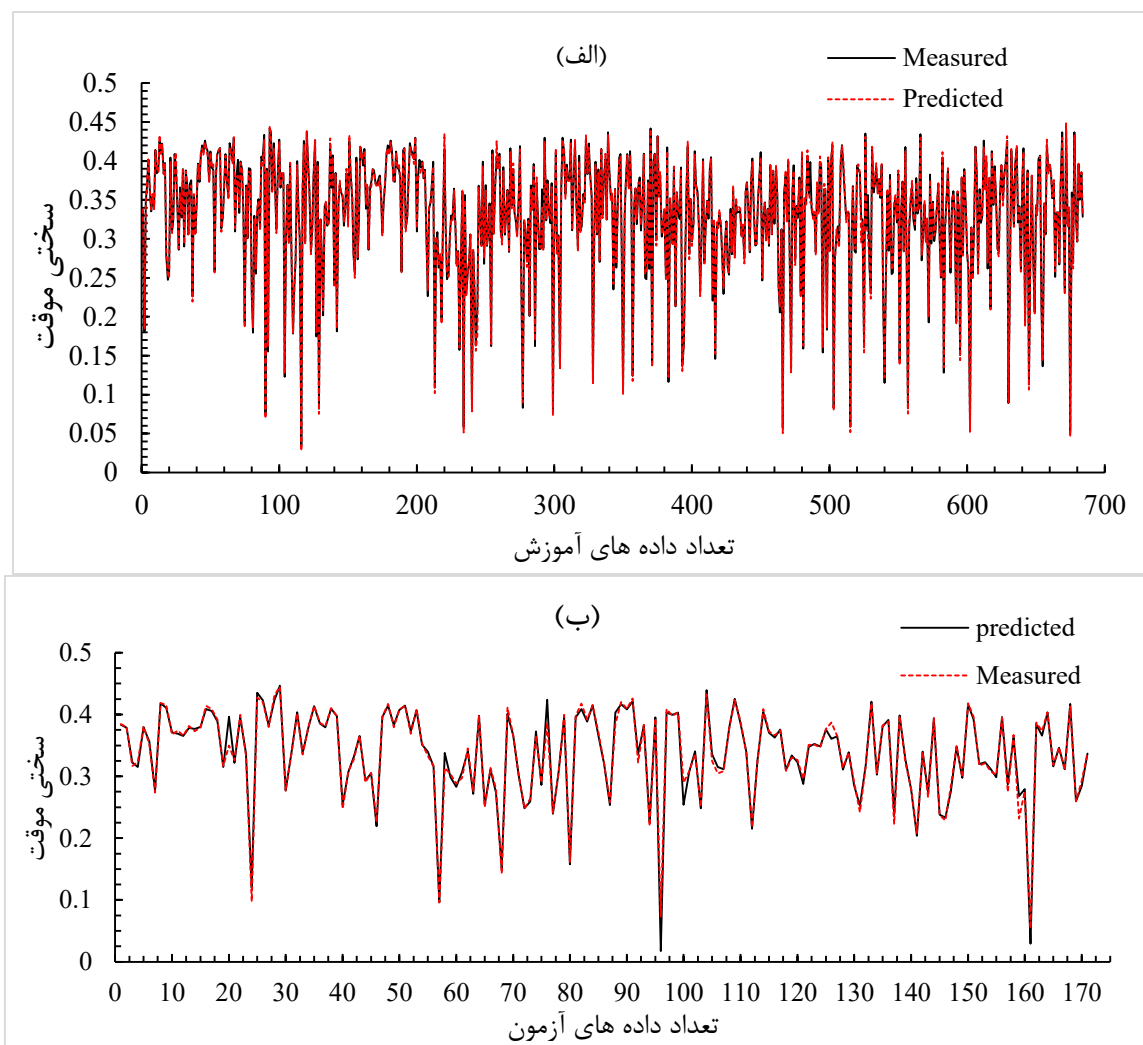
در ادامه نتایج بدست آمده از مدل‌سازی سختی موقت و دائم آب در مراحل آزمون و آموزش در مقایسه با داده‌های اندازه گیری شده ترسیم گردید. لازم به ذکر است نمودارهای ترسیم شده بر اساس داده‌های نرمالایز شده می‌باشد.



شکل ۴-۱۴۴. (الف) مقایسه سختی دائم پیش بینی شده و اندازه گیری در مرحله آموزش SVM

(ب) مقایسه سختی دائم پیش بینی شده و اندازه گیری در مرحله آزمون SVM

در شکل ۴-۱۴۴ مقایسه مقادیر اندازه گیری شده و تخمین زده شده سختی دائم آب در مرحله آزمون را نشان می دهد که نشان از تطابق مناسب پیش بینی و تغییرات سختی دائم در نمونه های آزمایش می باشد. اما ممکن است برخی از داده ها توسط ماشین بردار به درستی محاسبه نشده باشند که باعث ایجاد خطا می گردد.

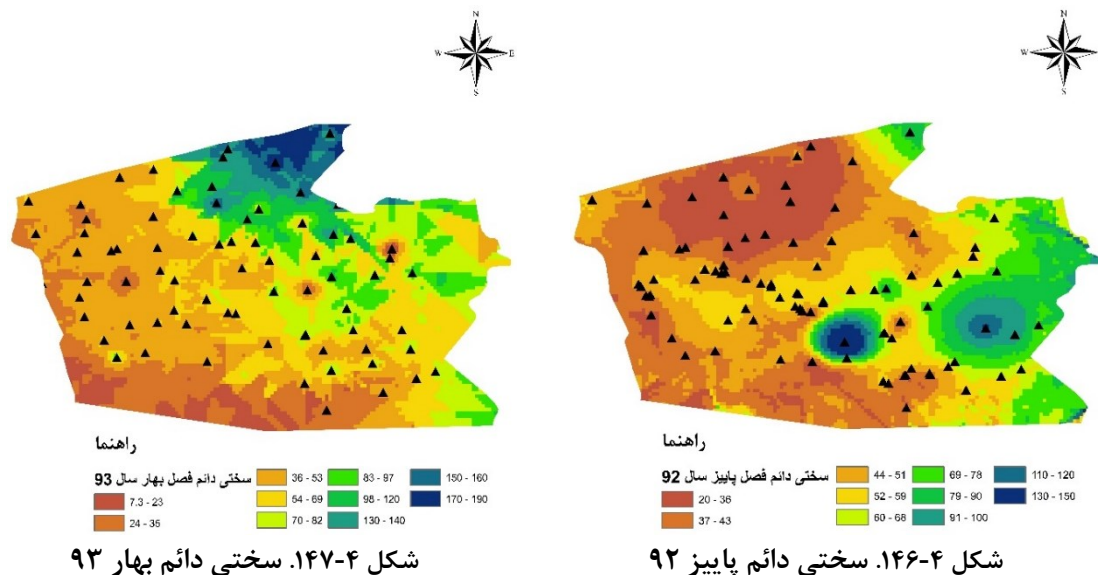


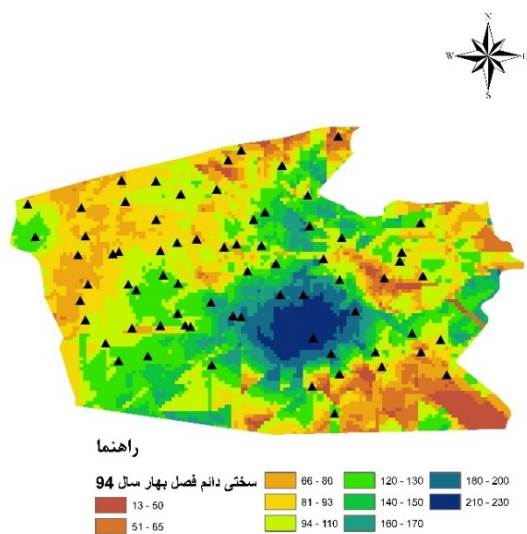
شکل ۴-۱۴۵. (الف) مقایسه سختی موقت پیش بینی شده و اندازه گیری در مرحله آموزش
(ب) مقایسه سختی موقت پیش بینی شده و اندازه گیری در مرحله آزمون

با توجه به شکل ۴-۱۴۵ (الف)، مدل ایجاد شده به روش ماشین بردار برای سختی دائم به ازای هر ورودی، خروجی آن دارای دقت مناسب بوده و نشان از مدلسازی بهینه می باشد. نتایج نشان داده شده در شکل ۴-۱۴۵ (ب) در مقایسه با نمودار ترسیم شده برای سختی دائم، دارای مقدار خطای بیشتری نسبت به آن است که در نمودار نیز مشهود است. در نهایت مدل های ایجاد شده با تقریب خوب قادر به پیش بینی غلظت سختی دائم و موقت می باشد. لازم به ذکر است نتایج نشان داده شده در نمودارهای فوق در حالت نرمالایز می باشد.

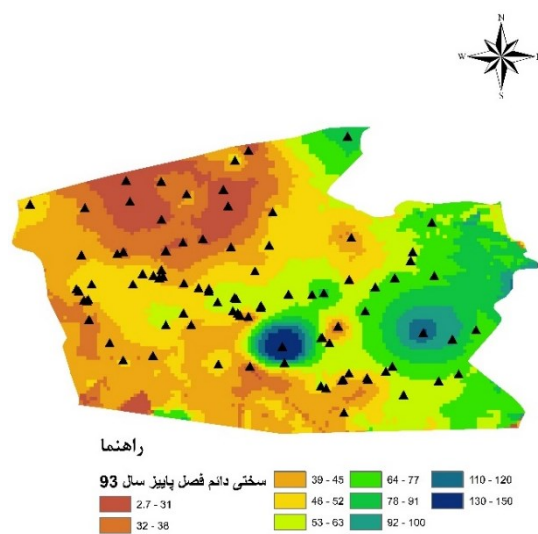
۱۳-۴ پیش‌بینی سختی موقت و دائم به کمک SVM

پس از ایجاد یک مدل بهینه برای تخمین سختی موقت و دائم با استفاده از داده‌های اندازه‌گیری شده در چاه‌های دشت بابل-آمل، اقدام به پیش‌بینی می‌نماییم. برای این منظور ابتدا با استفاده از ArcGIS تمام داده‌ها در کل محدوده مورد مطالعه را با استفاده از روش Kriging درونیابی کرده، سپس ویژگی هر نقطه را استخراج می‌کنیم. تعداد نقاط بدست آمده در هر نقشه، ۴۷۷۵۶۸ داده برای ۸ پارامتر در فصل بهار و پاییز سال ۹۲ تا ۹۸ تولید شد. داده‌های بدست آمده در بازه ۰ تا ۱ نرمالسازی گردید، سپس با قرار دادن داده‌های بدست آمده از ArcGIS، ماتریسی به ابعاد ۸×۴۷۷۵۶۸ برای ورود به مدل ماشین بردار آماده گردید. نتیجه بدست آمده از ورود داده‌ها به مدل تهیه شده، یک ماتریس به سطر یک و ستون ۴۷۷۵۶۸ می‌باشد که سختی موقت و دائم در فصل و سال مشخص شده است. نتایج با استفاده از کدنویسی در MATLAB از حالت نرمالسازی خارج و به مقیاس اصلی برگردانده شده است. شکل ۴-۱۴۶ تا ۴-۱۷۱ سختی دائم و موقت آب در محدوده مورد مطالعه در سال‌های ۹۲ تا ۹۸ نشان می‌دهد.

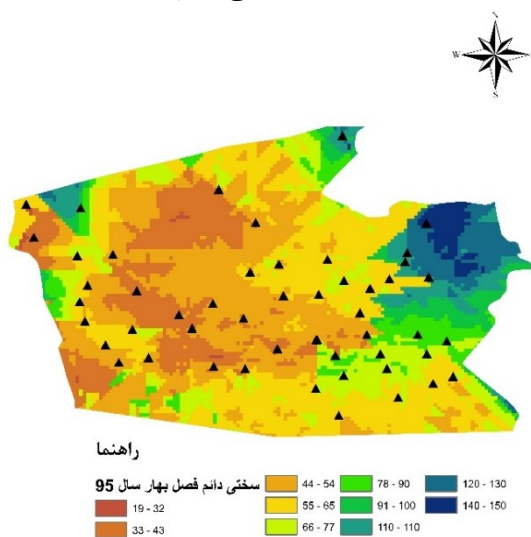




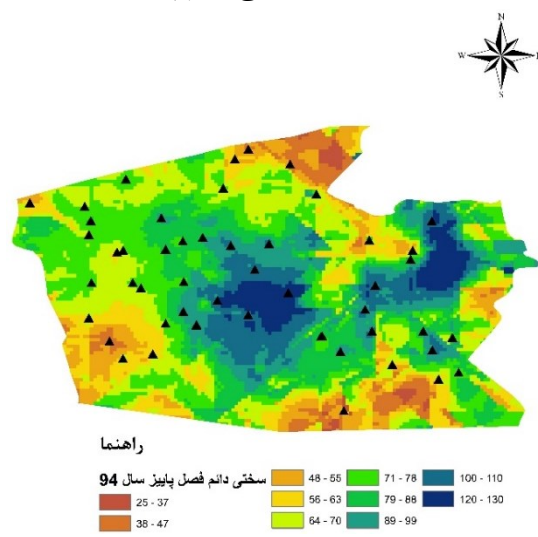
شکل ۴-۱۴۹. سختی دائم بهار ۹۴



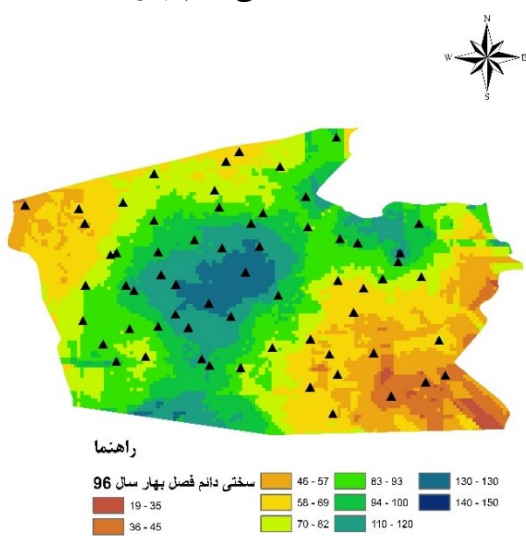
شکل ۴-۱۴۸. سختی دائم پاییز ۹۳



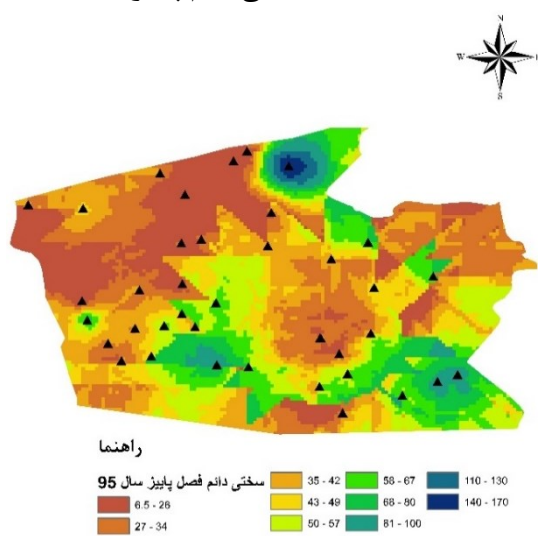
شکل ۴-۱۵۱. سختی دائم بهار ۹۵



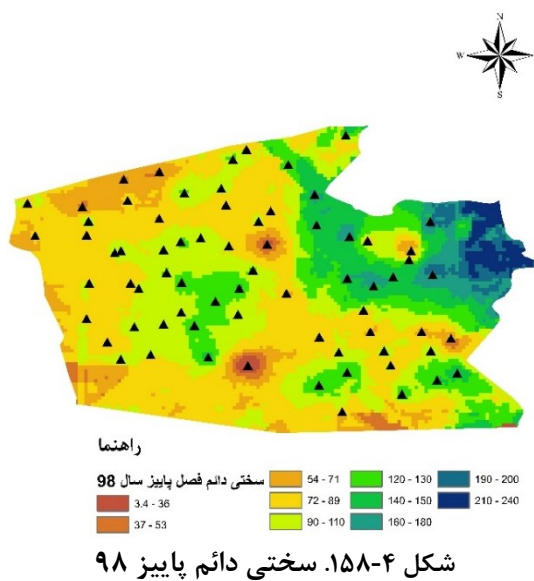
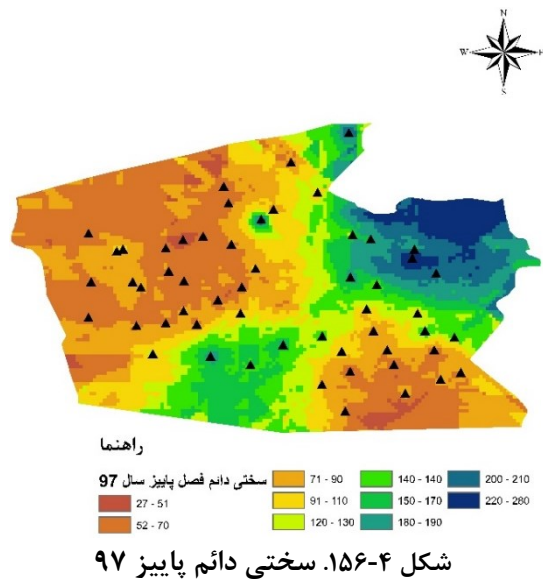
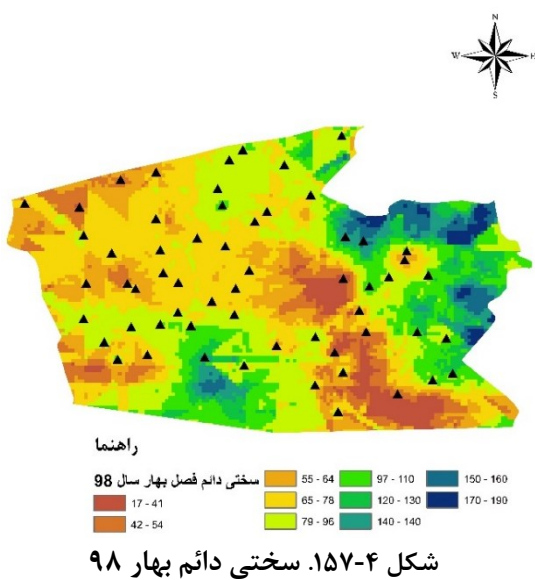
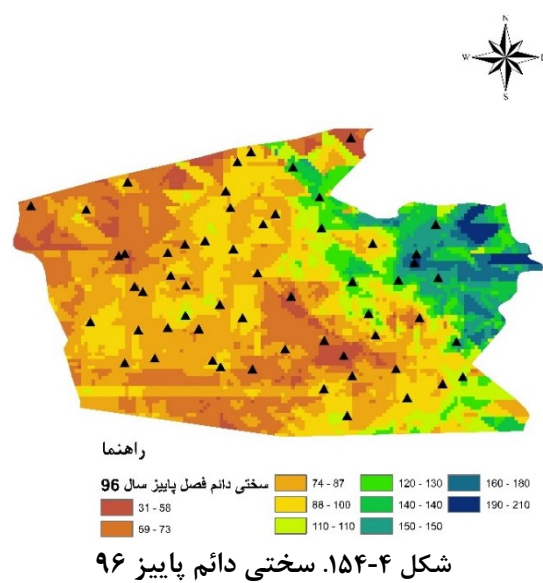
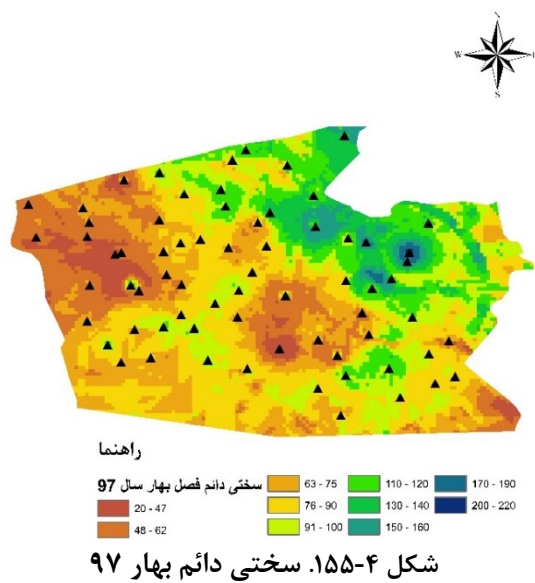
شکل ۴-۱۵۰. سختی دائم پاییز ۹۴



شکل ۴-۱۵۳. سختی دائم بهار ۹۶

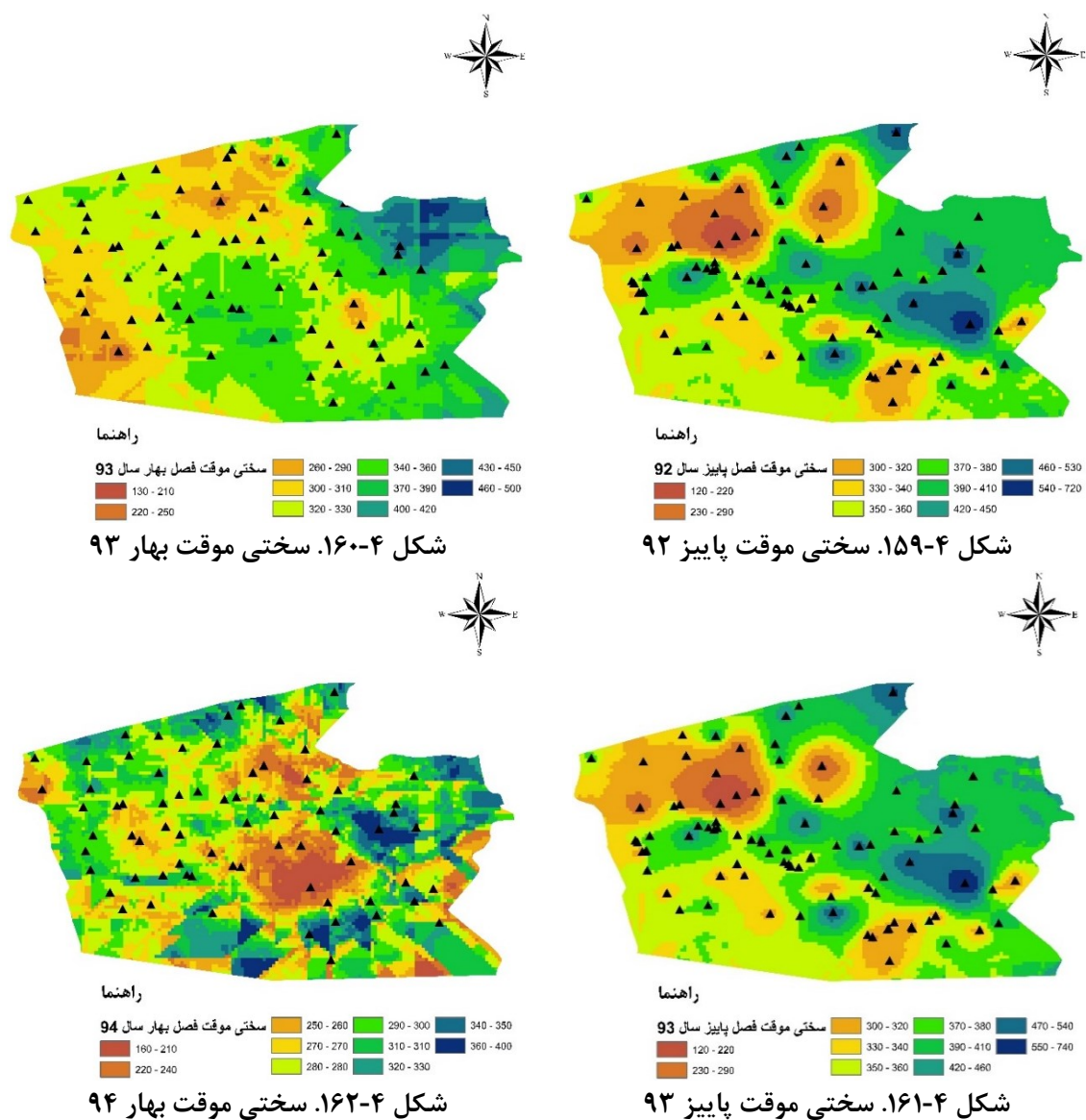


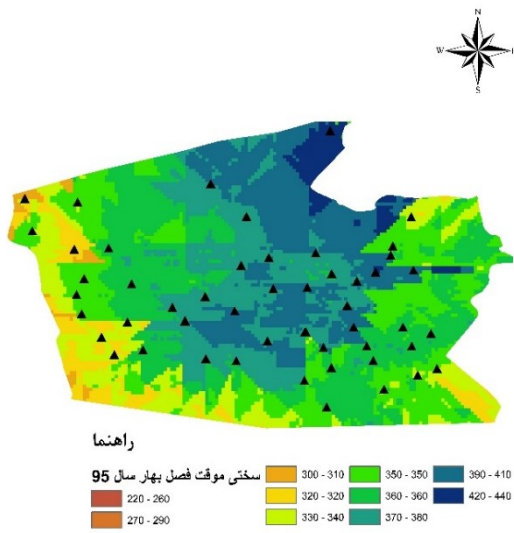
شکل ۴-۱۵۲. سختی دائم پاییز ۹۵



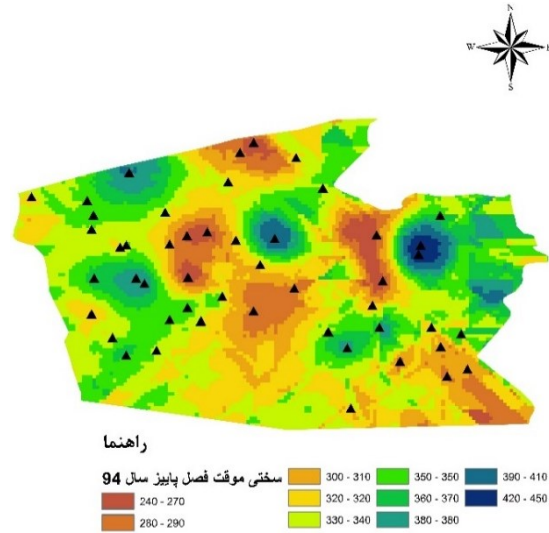
نتایج بدست آمده از روش SVM نشان می‌دهد، میزان سختی دائم بیشتر در فصل پاییز نسبت به فصل بهار بالاتر می‌باشد. علت آن را می‌توان به بارش‌هایی که در فصول بارندگی صورت می‌گیرد مرتبط دانست.

سختی موجود در آب توسط یون‌های فلزی چند ظرفیتی حاصل از سنگ‌های رسوبی، نشت و رواناب در خاک ایجاد می‌شود. دو یون اصلی کلسیم و منیزیم از ترکیبات معدنی ضروری مواد غذایی رایج هستند که در بسیاری از سنگ‌های رسوبی وجود داشته و عامل اصلی ایجاد سختی آب هستند.

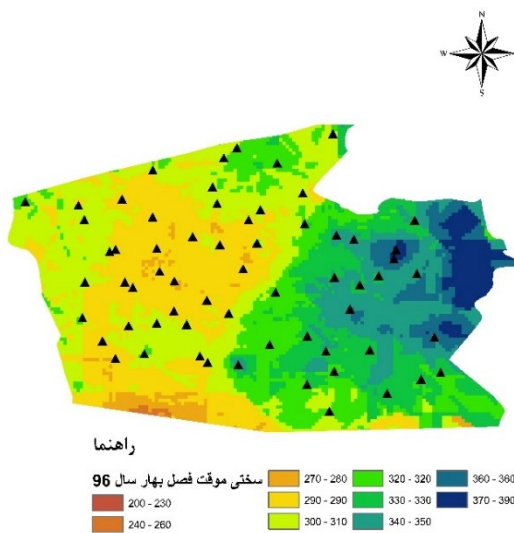




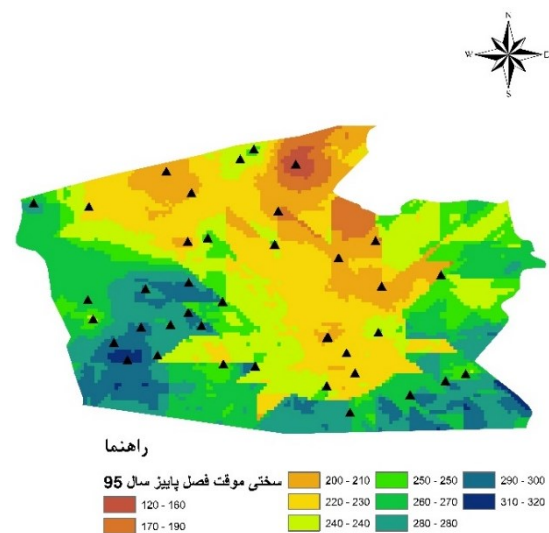
شکل ۴-۱۶۴. سختی موقت بهار ۹۵



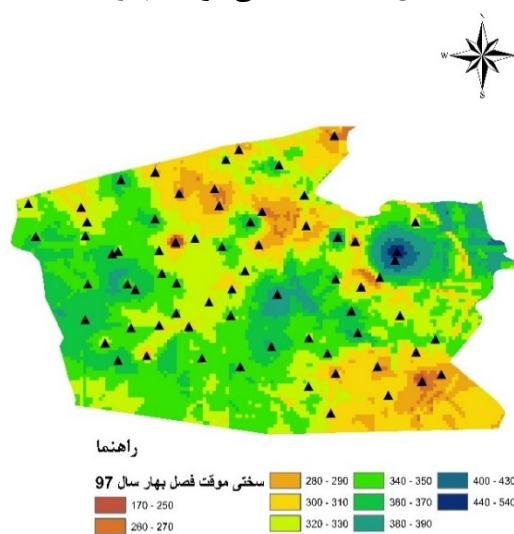
شکل ۴-۱۶۳. سختی موقت پاییز ۹۴



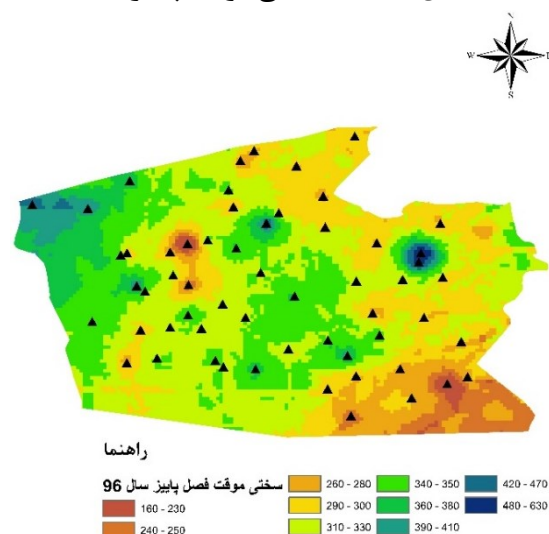
شکل ۴-۱۶۶. سختی موقت بهار ۹۶



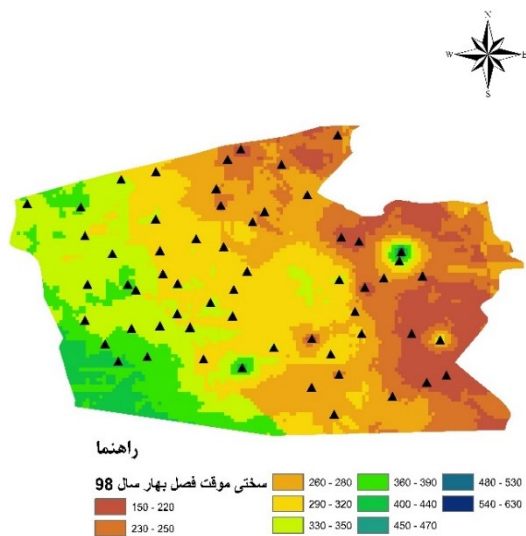
شکل ۴-۱۶۵. سختی موقت پاییز ۹۵



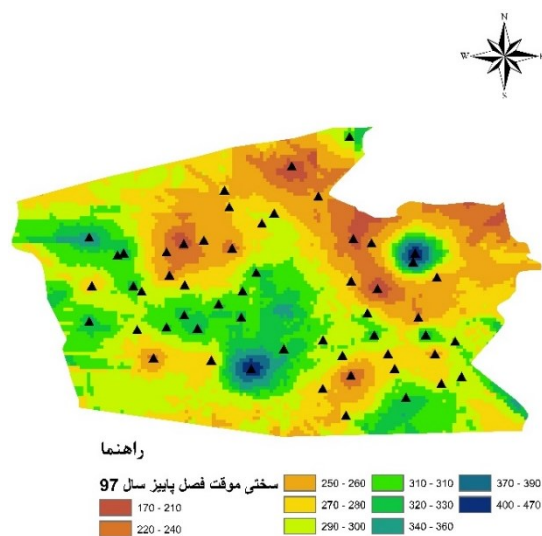
شکل ۴-۱۶۸. سختی موقت بهار ۹۷



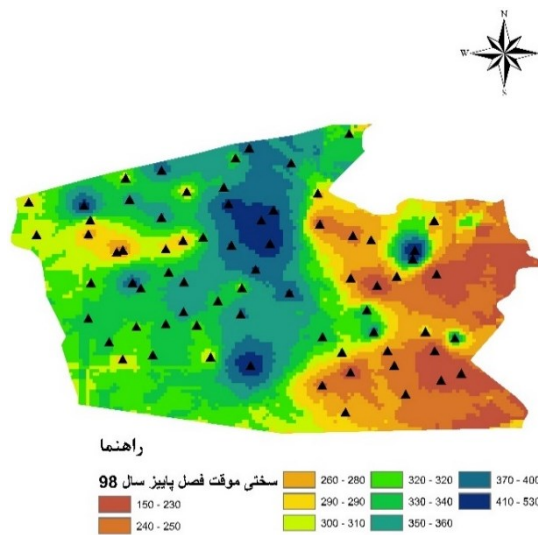
شکل ۴-۱۶۷. سختی موقت پاییز ۹۶



شکل ۴-۱۷۰. سختی موقت بهار ۹۸



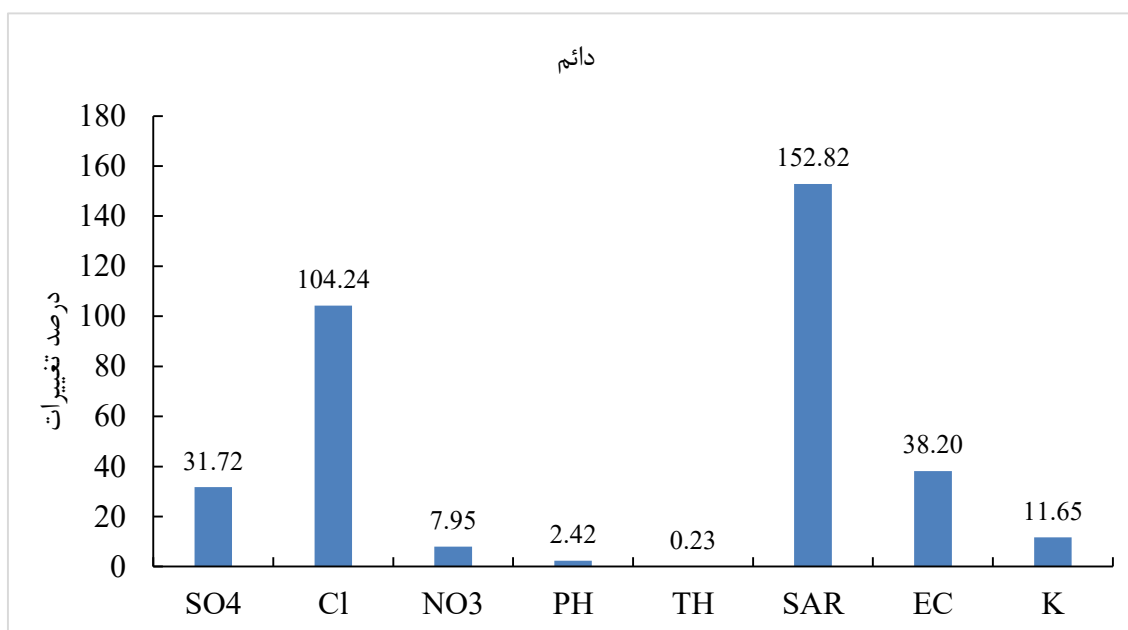
شکل ۴-۱۶۹. سختی موقت پاییز ۹۷

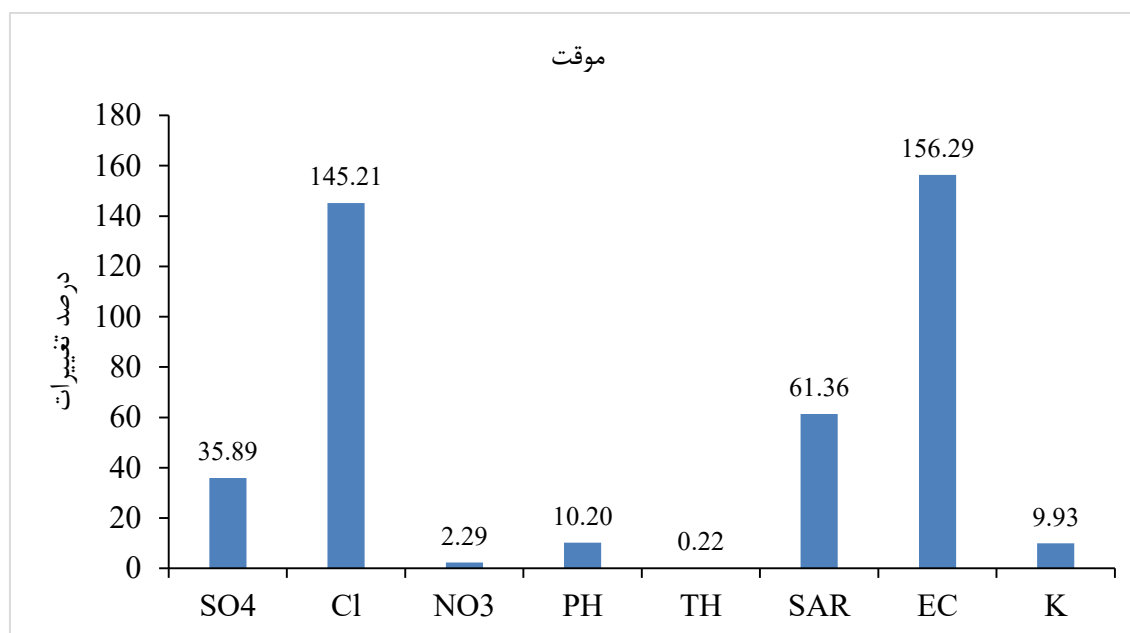


شکل ۴-۱۷۱. سختی موقت پاییز ۹۸

۱۴-۴ تحلیل حساسیت

تحلیل حساسیت ابزاری برای مطالعه تاثیر داده ها و پارامترها در نتیجه مدل مورد استفاده قرار می‌گیرد. این تحلیل برای یک یا چند پارامتر قابل انجام است. روش های مختلفی از جمله روش اتصال وزن که اولین بار توسط David Garson و سپس توسط Goh تحت عنوان تحلیل حساسیت برای بدست آوردن قواعد حاکم بر مدل استفاده شد. در این مطالعه، برای بررسی اثر هر پارامتر بر روی تخمین غلظت سختی دائم و موقت، برای هر پارامتر بصورت جداگانه یک مقدار ثابت (۵۰ درصد میزان اولیه غلظت) در شبکه آموزش داده شده وارد و حساسیت سختی دائم و موقت به هر پارامتر مورد ارزیابی قرار گرفت [۴۶].





شکل ۴-۱۷۲. نتایج بدست آمده از تحلیل حساسیت سختی دائم و موقت

نتایج بدست آمده از تغییرات به صورت قدر مطلق در شکل ۴-۱۷۱ و ۴-۱۷۲ نشان داده شد. همانطور که در شکل مشاهده می‌شود، پنج ورودی SO_4^{2-} ، Cl^- ، PH ، EC و K^+ بیشترین حساسیت را بر روی مدل خروجی در شبکه عصبی دارند. انتخاب داده‌های ورودی در مدلسازی های داده محور نقش کلیدی در رسیدن به یک مدل با کارایی مطلوب را دارد؛ بنابراین انتظار می‌رود با بکار بردن از داده‌هایی که بیشترین تاثیر را در تعیین مقادیر خروجی دارد مدلی با قابلیت اطمینان بالا ایجاد شود.

فصل ۵ : نتایج و مشاهدات

۵-۱ نتایج

رویکرد آماری مبتنی بر GIS امکان تخمین غلظت سطحی غلظت مواد آب زیرزمینی را که فقط از پارامترهای محیطی برداشت شده، می‌دهد. در این مطالعه با استفاده از پارامترهای PH ، Cl^- ، SO_4^{2-} ، NO_3^- و K^+ ، EC ، SAR ، TH جهت مدلسازی در شبکه عصبی پیشرو و ماشین بردار پشتیبان به منظور پیش بینی سختی موقت و دائم در آب‌های زیرزمینی پرداخته شد. برای این منظور، از داده‌های کیفی جمع آوری شده از ۸۵۵ چاه در دو فصل پاییز و زمستان از سال ۱۳۹۲ تا ۱۳۹۸ استفاده شده است. داده‌های ورودی برای آموزش از یک ماتریس 8×684 تشکیل شد. در نهایت نتایج بدست آمده با معیارهای ارزیابی (R^2 ، MSE و $RMSE$) بین مقادیر مشاهده شده

رویکرد آماری مبتنی بر GIS امکان تخمین غلظت سطحی غلظت مواد آب زیرزمینی را که فقط از پارامترهای محیطی برداشت شده، می‌دهد. در این مطالعه با استفاده از پارامترهای PH ، Cl^- ، SO_4^{2-} ، NO_3^- و K^+ ، EC ، SAR ، TH جهت مدلسازی در شبکه عصبی پیشرو و ماشین بردار پشتیبان به منظور پیش‌بینی سختی موقت و دائم در آب‌های زیرزمینی پرداخته شد. برای این منظور، از داده‌های کیفی جمع‌آوری شده از ۸۵۵ چاه در دو فصل پاییز و زمستان از سال ۱۳۹۲ تا ۱۳۹۸ استفاده شده است. داده‌های ورودی برای آموزش از یک ماتریس 8×684 تشکیل شد. در نهایت نتایج بدست آمده با معیارهای ارزیابی (R^2 ، MSE و $RMSE$) بین مقادیر مشاهده شده و پیش‌بینی شده، مورد تجزیه و تحلیل قرار گرفته شد. این معیار برای شبکه عصبی پیشرو برای سختی دائم به ترتیب 0.9303 ، $748/76$ و $27/36$ و برای سختی موقت برابر 0.9688 ، $504/78$ و $22/46$ می‌باشد. همچنین نتایج در ماشین بردار پشتیبان برای سختی دائم برابر 0.9304 ، 0.0001401 ، 0.0181 و برای سختی موقت برابر 0.9833 ، 0.0000338 ، 0.0058 بدست آمد. نتایج نشان می‌دهد، مدل ماشین بردار پشتیبان مناسب‌تری نسبت به شبکه عصبی خروجی دارد. علاوه بر این مدل شبکه عصبی در برخی نقاط توانایی محاسبه سختی را نداشته است. بطور کلی با ساخت مدل‌های مختلف در شبکه عصبی پیشرو و ماشین بردار پشتیبان با پارامترهای مختلف و مقایسه نتایج، تابعی ایجاد شد که تعمیم‌پذیری خوبی در خروجی دارد. این تابع با استفاده از داده‌های حقیقی بدست آمده از چاه‌ها آموزش دیده است. پس از ایجاد مدل بهینه، به تخمین غلظت سختی موقت و دائم در دشت آمل-بابل پرداخته شد. ورودی‌های مرتبط با مدل شبکه عصبی پیشرو و ماشین بردار پشتیبان برای پیش‌بینی به کمک درونیابی و با داشتن پارامترهای مربوط به چاه‌های مورد نظر در فصل پاییز و زمستان استخراج و سپس به کمک آن مقادیر سختی موقت و دائم در منطقه مطالعاتی پیش‌بینی گردید.

نتایج حاصل از مدلسازی نشان دهنده توانایی مدل‌های ایجاد شده در تعیین سختی در آب‌های زیرزمینی در مقایسه با سایر مدل‌هایی که نیاز به پارامترهای هیدرولوژیکی دارند را نشان می‌دهد.

غلظت سختی در قسمت شرق و مرکزی نسبت به سایر نقاط بیشتر بوده بطوری که از سمت شرق به مرکز حرکت می کند.

۲-۵ پیشنهادات

باتوجه به آنچه که در این تحقیق انجام شد موارد زیر برای انجام تحقیقات آتی می تواند مورد استفاده قرار گیرد:

۱- برای حصول نتیجه دقیق تر می توان از داده های هواشناسی و سموم کشاورزی بعنوان پارامترهای ورودی مدل استفاده نمود.

۲- استفاده از الگوریتم های فرا ابتکاری جهت بهینه سازی پارامترهای مدلسازی

۳- تخمین غلظت سختی در دشت آمل- بابل به کمک سایر ابزارهای هوش مصنوعی قوی تر با

معیارهای خطای کمتر مانند الگوریتم XGboost در یادگیری ماشین

۴- به منظور بررسی دقیق تر، می توان کاربری اراضی را در نظر گرفت

۵- برای بررسی اطلاعات در سال های آینده می توان از سری های زمانی استفاده نمود.

- [1] امین علیزاده، اصول هیدرولوژی کاربردی. دانشگاه امام رضا (ع)، ۱۳۸۲.
- [2] Y. Takahashi and Y. Imaizumi, "Hardness in Drinking Water," *Eisei kagaku*, vol. 34, no. 5, pp. 475–479, 1988, doi: 10.1248/jhs1956.34.475.
- [3] A. J. B. Zehnder, Ed., *Soil and Groundwater Pollution*. Dordrecht: Springer Netherlands, 1995.
- [4] S. Akram, "Hardness in Drinking-Water, its Sources, its Effects on Humans and its Household Treatment," *J. Chem. Appl.*, vol. 4, no. 1, pp. 01–04, 2018, doi: 10.13188/2380-5021.1000009.
- [5] R. S. Govindaraju and A. R. Rao, Eds., *Artificial Neural Networks in Hydrology*, vol. 36. Dordrecht: Springer Netherlands, 2000.
- [6] ASCE Task Committee on Application of Artificial Neural Network in Hydrology, "Artificial Neural Network in Hydrology. I: Preliminary Concepts," *J. Hydrol. Eng.*, vol. 5, no. 2, pp. 115–123, 2000.
- [7] Ke-Lin Du and M. N. S. Swamy, *Neural Networks in a Softcomputing Framework*, 1st editio. London: Springer-Verlag, 2006.
- [8] D. P. Mandic and J. A. Chambers, *Recurrent Neural Networks for Prediction: Learning algorithms, architectures and stability.*, vol. 4, no. 2. 2001.
- [9] Q. J. Zhang and K. C. Gupta, *Neural Networks for RF and Microwave Design*. 2000.
- [10] K. Abhishek, A. Kumar, R. Ranjan, and S. Kumar, "A rainfall prediction model using artificial neural network," in *2012 IEEE Control and System Graduate Research Colloquium*, Jul. 2012, pp. 82–87, doi: 10.1109/ICSGRC.2012.6287140.
- [11] R. Barzegar, A. Asghari Moghaddam, J. Adamowski, and E. Fijani, "Comparison of machine learning models for predicting fluoride contamination in groundwater," *Stoch. Environ. Res. Risk Assess.*, vol. 31, no. 10, pp. 2705–2718, Dec. 2017, doi: 10.1007/s00477-016-1338-z.
- [12] M. Isazadeh, S. M. Biazar, and A. Ashrafzadeh, "Support vector machines and feed-forward neural networks for spatial modeling of groundwater qualitative parameters," *Environ. Earth Sci.*, vol. 76, no. 17, p. 610, Sep. 2017, doi: 10.1007/s12665-017-6938-5.
- [13] J. P. Davim, Ed., *Machining of Hard Materials*. London: Springer London, 2011.
- [14] L. Wang, Y. Zeng, and T. Chen, "Back propagation neural network with adaptive differential evolution algorithm for time series forecasting," *Expert Syst. Appl.*, vol. 42, no. 2, pp. 855–863, Feb. 2015, doi: 10.1016/j.eswa.2014.08.018.
- [15] S. Sharma, S. Sharma, and A. Athaiya, "Activation Functions in Neural Networks," *Int. J. Eng. Appl. Sci. Technol.*, vol. 04, no. 12, pp. 310–316, 2020, doi: 10.33564/ijeast.2020.v04i12.054.
- [16] R. G. Brereton and G. R. Lloyd, "Support Vector Machines for classification and

- regression,” *Analyst*, vol. 135, no. 2, pp. 230–267, 2010, doi: 10.1039/b918972f.
- [17] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY: Springer New York, 2000.
- [18] V. N. Vapnik, *Statistical Learning Theory*, 1st ed. Wiley, 1998.
- [19] “Introduction to Support Vector Machines.”
https://docs.opencv.org/4.5.2/d1/d73/tutorial_introduction_to_svm.html.
- [20] Q. Quan, Z. Hao, H. Xifeng, and L. Jingchun, “Research on water temperature prediction based on improved support vector regression,” *Neural Comput. Appl.*, Mar. 2020, doi: 10.1007/s00521-020-04836-4.
- [21] A. Hipni, A. El-shafie, A. Najah, O. A. Karim, A. Hussain, and M. Mukhlisin, “Daily Forecasting of Dam Water Levels: Comparing a Support Vector Machine (SVM) Model With Adaptive Neuro Fuzzy Inference System (ANFIS),” *Water Resour. Manag.*, vol. 27, no. 10, pp. 3803–3823, Aug. 2013, doi: 10.1007/s11269-013-0382-4.
- [22] A. J. SMOLA and B. SCHOLKOPF, “A tutorial on support vector regression,” *Stat. Comput.*, vol. 14, pp. 199–222, 2004, [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=1CAD92EF8CCE726A305D8A41F873EEFC?doi=10.1.1.114.4288&rep=rep1&type=pdf%0Ahttp://download.springer.com/static/pdf/493/art%3A10.1023%2FB%3ASTCO.0000035301.49549.88.pdf?auth66=1408162706_8a28764ed0fae9.
- [23] ر. عربگل، “بررسی کارایی مدل ماشینهای بردار پشتیبان در پیش بینی غلظت نیترات در آبهای زیرزمینی و مقایسه آن با مدل ”modflow صنعتی اصفهان، ۱۳۹۰.
- [24] S. Winters-Hilt and S. Merat, “SVM clustering,” *BMC Bioinformatics*, vol. 8, no. S7, p. S18, Nov. 2007, doi: 10.1186/1471-2105-8-S7-S18.
- [25] World Health Organisation, “Hardness in drinking-water. Background document for development,” *World Heal. Organ.*, 2009.
- [26] J. C. R. R. T. D. W. H. K. J. H. G. T. Crittenden, *MWH’s Water Treatment: Principles and Design*, 3rd editio. John Wiley & Sons, 2012.
- [27] M. A. O. X. B, S. Songhao, and L. I. U. Xiang, “Groundwater Level Predictions Using Artificial Neural Networks,” vol. 7, no. 6, pp. 574–579, 2002.
- [28] G. B. Sahoo, C. Ray, E. Mehnert, and D. A. Keefer, “Application of artificial neural networks to assess pesticide contamination in shallow groundwater,” *Sci. Total Environ.*, vol. 367, no. 1, pp. 234–251, Aug. 2006, doi: 10.1016/j.scitotenv.2005.12.011.
- [29] A. Gemitzi, C. Petalas, V. Pisinaras, and V. A. Tsihrintzis, “Spatial prediction of nitrate pollution in groundwaters using neural networks and GIS: an application to South Rhodope aquifer (Thrace, Greece),” *Hydrol. Process.*, vol. 23, no. 3, pp. 372–383, Jan. 2009, doi: 10.1002/hyp.7143.
- [30] K. Al-Mahallawi, J. Mania, A. Hani, and I. Shahrour, “Using of neural networks for the prediction of nitrate groundwater contamination in rural and agricultural areas,” *Environ. Earth Sci.*, vol. 65, no. 3, pp. 917–928, Feb. 2012, doi: 10.1007/s12665-011-1134-5.

- [31] M. M. Hossain and M. Piantanakulchai, "Groundwater arsenic contamination risk prediction using GIS and classification tree method," *Eng. Geol.*, vol. 156, pp. 37–45, Apr. 2013, doi: 10.1016/j.enggeo.2013.01.007.
- [32] V. Rodriguez-Galiano, M. P. Mendes, M. J. Garcia-Soldado, M. Chica-Olmo, and L. Ribeiro, "Predictive modeling of groundwater nitrate pollution using Random Forest and multisource variables related to intrinsic and specific vulnerability: A case study in an agricultural setting (Southern Spain)," *Sci. Total Environ.*, vol. 476–477, pp. 189–206, Apr. 2014, doi: 10.1016/j.scitotenv.2014.01.001.
- [33] R. Arabgol, M. Sartaj, and K. Asghari, "Predicting Nitrate Concentration and Its Spatial Distribution in Groundwater Resources Using Support Vector Machines (SVMs) Model," *Environ. Model. Assess.*, vol. 21, no. 1, pp. 71–82, Jan. 2016, doi: 10.1007/s10666-015-9468-0.
- [34] K. Ostad-Ali-Askari, M. Shayannejad, and H. Ghorbanizadeh-Kharazi, "Artificial neural network for modeling nitrate pollution of groundwater in marginal area of Zayandeh-rood River, Isfahan, Iran," *KSCE J. Civ. Eng.*, vol. 21, no. 1, pp. 134–140, Jan. 2017, doi: 10.1007/s12205-016-0572-8.
- [35] H. Darwishe, J. El Khattabi, F. Chaaban, B. Louche, E. Masson, and E. Carlier, "Prediction and control of nitrate concentrations in groundwater by implementing a model based on GIS and artificial neural networks (ANN)," *Environ. Earth Sci.*, vol. 76, no. 19, p. 649, Oct. 2017, doi: 10.1007/s12665-017-6990-1.
- [36] T. Sheikhy Narany, M. F. Ramli, A. Z. Aris, W. N. A. Sulaiman, and K. Fakharian, "Groundwater irrigation quality mapping using geostatistical techniques in Amol–Babol Plain, Iran," *Arab. J. Geosci.*, vol. 8, no. 2, pp. 961–976, Feb. 2015, doi: 10.1007/s12517-014-1271-8.
- [37] T. S. Narany, M. F. Ramli, K. Fakharian, and A. Z. Aris, "A GIS-index integration approach to groundwater suitability zoning for irrigation purposes," *Arab. J. Geosci.*, vol. 9, no. 7, p. 502, Jun. 2016, doi: 10.1007/s12517-016-2520-9.
- [38] F. Edition, *Guidelines for Drinking-water Quality*, FOURTH EDI. 2011.
- [39] C. E. Boyd, *Water Quality*. Cham: Springer International Publishing, 2020.
- [40] "Conductivity," *EPA*, 2012.
<https://archive.epa.gov/water/archive/web/html/vms59.html#:~:text=Conductivity in water is affected,that carry a positive charge>.
- [41] E. Cotruvo J, Bartram J, "Calcium and Magnesium in Drinking-water," *World Heal. Organ.*, pp. 1–194, 2009, [Online]. Available: http://apps.who.int/iris/bitstream/handle/10665/43836/9789241563550_eng.pdf;jsessionid=D351ABBE791C3A84AAF8C2BD7DD12C2D?sequence=1.
- [42] G. S. Bhunia, P. K. Shit, and R. Maiti, "Comparison of GIS-based interpolation methods for spatial distribution of soil organic carbon (SOC)," *J. Saudi Soc. Agric. Sci.*, vol. 17, no. 2, pp. 114–126, Apr. 2018, doi: 10.1016/j.jssas.2016.02.001.
- [43] A. Mirarabi, H. R. Nassery, M. Nakhaei, J. Adamowski, A. H. Akbarzadeh, and F. Alijani, "Evaluation of data-driven models (SVR and ANN) for groundwater-

- level prediction in confined and unconfined systems,” *Environ. Earth Sci.*, vol. 78, no. 15, p. 489, Aug. 2019, doi: 10.1007/s12665-019-8474-y.
- [44] H. Nguyen, “Support vector regression approach with different kernel functions for predicting blast-induced ground vibration: a case study in an open-pit coal mine of Vietnam,” *SN Appl. Sci.*, vol. 1, no. 4, p. 283, Apr. 2019, doi: 10.1007/s42452-019-0295-9.
- [45] J. Han, *Data Mining: Concepts and Techniques (The Morgan Kaufmann Series in Data Management Systems)*. 2011.
- [46] F. Qaderi and E. Babanezhad, “Prediction of the groundwater remediation costs for drinking use based on quality of water resource, using artificial neural network,” *J. Clean. Prod.*, vol. 161, pp. 840–849, Sep. 2017, doi: 10.1016/j.jclepro.2017.05.187.

Abstract

Whenever there are no other resources to provide water, groundwater resources are the only source in many regions. Reduction in groundwater, pollution and improper management are some issues that affected human life. Considering the role of water in human health, its quality must be inspected via conducting researches for quality control. In this area, smart and data-centered models are rapidly developing methods due to their ease of use compared with others. These models can learn and generalize data and are able to solve problems of prediction, management, and estimation in various water management aspects. Meanwhile, neural networks are powerful tools in solving regression and classification problems that can solve problems with complex hydrology and low data compared to data-centered methods. In this study, a feedforward neural network method and support vector machine (SVM) method is employed to predict temporary and permanent hardness concentrations of Amol-Babol plain using chemical parameters of water, such as SO_4^{2-} , Cl^- , NO_3^- , K^+ , EC , SAR , TH and PH . For this purpose, the qualitative data collected from 855 wells during two fall and spring seasons in six years was used. Then, by creating a feedforward neural network method and SVM model with different parameters, the optimal model was selected, trained, and tested. Based on R^2 , MSE and RMSE criteria, the obtained results were 0.9303, 748.76, and 27.36, respectively, for feedforward neural network at the permanent hardness and 0.9688, 504.78, and 22.46, respectively, for temporary hardness. These results are, respectively, 0.9304, 0.0001401, and 0.0181 for permanent hardness and 0.9833, 0.0000338 and 0.0058, respectively, for temporary hardness in the SVM.

The obtained results of temporary and permanent hardness in Amol-Babol plain by data-centred models showed the capability of created models in modelling, which predicted hardness concentration in the desired area with an acceptable approximation.

Keywords: Support Vector Machine, Neural Network, GroundWater, Permanent Hardness, Temporary Hardness



Shahrood University of
Technology

Faculty of Civil Engineering

M.Sc.(M.A. or Ph.D) Thesis in Water Resource and Hydraulic Structure

**Prediction of temporary and permanent hardness Concentration in
Groundwater based on Chemistry Parameters by Combining Neural
Network and Geographic Information System (GIS)**

By: S.Mostafa Mousavi

Supervisor:
Dr. Ahmad Ahmadi
Dr. Farhad Qaderi

september,2021