

مقایسه الگوریتم تنبیل و مدل M5 در برآورد تراز سطح ایستابی (مطالعه موردی دشت نیشابور)

آتنا خلیلی نفت چالی^{۱*}، علی شهیدی^۲، عباس خاشعی سیوکی^۳

atenakhalili_2014@yahoo.com

دانشجوی دکتری مهندسی علوم آب دانشگاه بیرجند (۰۹۳۰۶۵۲۰۱۱۶)

Ashahidi@birjand.ac.ir

دانشیار گروه مهندسی علوم آب دانشگاه بیرجند

abbaskhashei@ birjand.ac.ir

استادیار گروه مهندسی علوم آب دانشگاه بیرجند

Atena khalili naft chali^{۱*}، ali shahidi^۲، abbas khashei siuki^۳

1- PhD student of water Science engineering، University of Birjand* Corresponding Author
(09306520116، Email: Atenakhalili_2014@yahoo.com)

2 - Associate professor، Department of water Science engineering، University of Birjand

3- Assistant professor، Department of water Science engineering، University of Birjand

چکیده:

در سال‌های اخیر در بسیاری از کشورها برداشت آب منابع زیرزمینی از میزان تغذیه سالیانه آن‌ها بیشتر بوده است. این امر باعث پایین افتادن سطح آب زیرزمینی و به دنبال آن خشک شدن چاه‌ها، قنات‌ها و چشمه شده است. در این تحقیق با توجه به اهمیت دشت نیشابور در تأمین آب کشاورزی، شرب و صنعت منطقه، از الگوریتم‌های تنبیل KSTAR، KNN و LWL و مدل درختی M5 تحت هفت سناریوی مختلف، برای برآورد تراز سطح ایستابی این آبخوان استفاده شده است. برای مقایسه نتایج، آماره‌های ریشه متوسط خطای مربعات، ضریب همبستگی و متوسط قدر مطلق خطا مورد تحلیل قرار گرفت. نتایج نشان داد که سناریو f که شامل پارامترهای حجم آب تخلیه شده و مجموع بارندگی می‌باشد، به دلیل در نظر نگرفتن پارامتر تراز سطح زمین، کارایی کمتری دارد. در سناریوی a، b و g در نظر گرفتن پارامترهای مجموع بارندگی در ماه قبل، مجموع بارندگی در دو ماه قبل و تراز سطح زمین، برآورد مطلوب‌تری از تراز سطح ایستابی حاصل می‌شود. در بین مدل‌های الگوریتم تنبیل و مدل درخت تصمیم M5 توانایی مدل KNN تحت سناریوی a در ماه آذر با داشتن $R^2 = 0.96$ ، $RMSE = 6/56$ و $MAE = 3/53$ از سایر مدل‌ها بیشتر می‌باشد. همچنین بررسی معیارهای ارزیابی نشان داد که مدل LWL، مدل مناسبی برای پیش بینی تراز سطح ایستابی نمی‌باشد.

واژه کلیدی: الگوریتم تنبیل، درخت تصمیم M5، تراز سطح ایستابی، دشت نیشابور

مقدمه:

آب‌های زیرزمینی به‌عنوان تنها منبع مورد اعتماد مصرف آب در زمینه‌های شرب، کشاورزی و صنعت در مناطق خشک و نیمه‌خشک محسوب می‌شوند. با توجه به اینکه ریزش‌های جوی در این مناطق عمدتاً به‌گونه‌ای است که در زمان‌های محدود اتفاق می‌افتد، لذا جریان‌های سطحی عموماً به‌صورت سیلابی و پس از هر بارش مشاهده و قطع می‌گردد. بنابراین در این مناطق مهار و بهره‌برداری از آب‌های سطحی فقط به شکل پروژه‌های آبخیزداری و در جهت تغذیه آب‌های زیرزمینی انجام می‌گیرد، لذا کلیه برنامه‌ریزی‌های کوتاه‌مدت و درازمدت تأمین آب در این مناطق، بر مبنای حجم آب زیرزمینی در دسترس است (۱۲). نوسان تراز سطح ایستابی در طول زمان باعث ایجاد جریان غیر ماندگار می‌شود و به تبع آن دبی خروجی نسبت به زمان متغیر خواهد بود (۱۷). مدل کردن سفره آب‌های زیرزمینی، به‌منظور پیش‌بینی سطح ایستابی، از نظر ایجاد سازه‌های مهندسی، مطالعات هیدرولوژی و مدیریتی، مصارف کشاورزی و به دست آوردن آب‌های زیرزمینی با کیفیت بالا، از اهمیت بالایی برخوردار است. در بسیاری از موارد تغییرات سطح ایستابی در آبخوان‌ها صدمات جبران‌ناپذیری را به سازه‌های مهندسی وارد می‌کند. با پی بردن به این تغییرات می‌توان از نظر بیلان هیدروژئولوژی و اهداف مدیریتی، تصمیمات مناسب‌تری را ارائه نمود و آب‌های زیرزمینی با کیفیت بالاتری را بدست آورد (۸).

تاکنون مدل‌های زیادی جهت پیش‌بینی سطح آب زیرزمینی به‌کاربرده شده است. از جمله این مدل‌ها می‌توان به مدل‌های سری زمانی تجربی و مدل‌های فیزیکی اشاره کرد (۲). فروغی و همکاران (۱۹) یک مدل شبکه‌ی عصبی را برای پیش‌بینی نوسان‌های سطح ایستابی با استفاده از اطلاعات بارش، تبخیر، اختلاف دبی در طول رودخانه و میزان برداشت ماهانه از منابع آب زیرزمینی برای حوضه‌ی آبریز کارون در استان خوزستان ارائه کردند. مختاری و همکاران (۱۳) در تحقیقی از مدل ترکیبی کریجینگ عصبی به‌عنوان روشی جدید برای پیش‌بینی زمانی و مکانی تراز آب‌های زیرزمینی در آبخوان محدوده دشت شبستر استفاده کردند. با توجه به مدل کریجینگ عصبی، بهترین مدل ارائه‌شده مربوط به مدل FNN-LM در ترکیب با مدل کریجینگ بلوکی بود. ابارشی و همکاران (۱) در پژوهشی به مقایسه عملکرد سه روش هوش مصنوعی برای پیش‌بینی نوسانات تراز سطح ایستابی پرداخته‌اند. نتایج نشان داد که هر سه روش شبکه‌های عصبی، فازی-عصبی و برنامه‌ریزی بیان ژن می‌توانند به‌طور موفقیت‌آمیزی برای پیش‌بینی نوسانات سطح ایستابی مورد استفاده قرار گیرند. از طرفی عملکرد مدل برنامه‌ریزی بیان ژن در مقایسه با دو روش دیگر از جنبه‌های مختلف بهتر بود. در روش برنامه‌ریزی بیان ژن می‌توان رابطه ریاضی حاکم بر مسئله مورد بررسی را به‌صورت

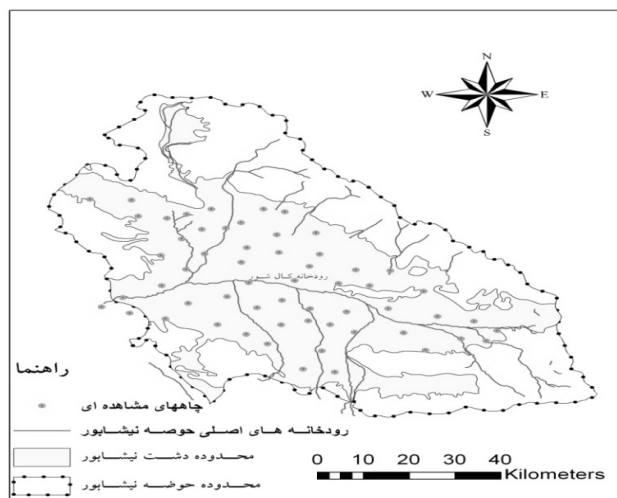
دقیق مشخص نمود. خلیلی پور (۶) در بررسی روند کمی و کیفی آب‌های زیرزمینی دشت قم نشان داده که متوسط سطح ایستابی در طی سال‌های ۱۳۶۸ الی ۱۳۷۹، حدود ۱۴/۵ متر افت داشته است و حداکثر افت سالانه در منطقه رقمی معادل ۴/۵ متر در سال بوده است. سبتی (۲۴) در سال ۱۹۸۹ با فرض اینکه سطح اولیه ایستابی به شکل سهمی است، یک روش تحلیلی و یک روش عددی برای پیش‌بینی جریان ناپایدار سطح ایستابی برای سیستم زهکشی در دو عمق مختلف ارائه نمود و نشان داد که دو روش فوق برای خاک‌های یکنواخت نتایج مشابهی ارائه می‌دهند. خاشعی سیوکی و همکاران (۵) به مقایسه مدل‌های شبکه عصبی مصنوعی، ANFIS و رگرسیونی در برآورد سطح ایستابی آبخوان دشت نیشابور پرداختند. نتایج روش‌های مذکور نشان داد که برای مدل ANN بهترین دقت به ورودی a در آذر ماه نسبت داده شد و در مدل ANFIS ورودی b در شهریورماه بهترین کارایی را داشتند. مدل رگرسیونی همواره کمترین دقت را داشته‌اند. زارع آبیانه و همکاران (۷) سطح ایستابی دشت ملایر بر اساس داده‌های هواشناسی با استفاده از شبکه‌ی عصبی مصنوعی شبیه‌سازی کردند. نتایج نشان داد که به کارگیری عوامل سطح ایستابی سال‌های گذشته، در ورودی شبکه‌ی عصبی بسیار حائز اهمیت است. با توجه به دقت مناسب مدل و روند کاهنده‌ی حاکم بر سفره، می‌توان استفاده از شبکه‌ی عصبی مصنوعی برای تصمیم‌گیری در مدیریت دشت ملایر توصیه کرد. حمزه و همکاران (۴) تغییرات پروفیل سطح ایستابی و دبی خروجی از زهکش‌های پلکانی را در یک خاک مطابق بررسی کردند.

روش‌های داده‌کاوی، روش‌های مدل کردن رابطه نهفته در داده‌ها هستند که به صورت خودکار به دسته‌بندی مجموعه داده‌ها (معمولاً مجموعه‌های بزرگ) و کشف ارتباط نهفته در بین آن‌ها به منظور قابل فهم شدن و در نتیجه سودمند شدن آن‌ها می‌پردازند (۲۰). پرکاربردترین این مدل‌ها عبارتند از: شبکه‌های عصبی مصنوعی، درخت تصمیم، سیستم‌های استنتاج فازی و ماشین‌های بردار پشتیبان. درختان تصمیم ابزاری است که قابلیت پاسخ‌گویی به مسائل پیچیده و غیرخطی را دارد و از آنجایی که در علوم دیگر از جمله برق و الکترونیک که اساساً شامل مسائل غیرخطی هستند به‌خوبی از عهده حل مسائل برآمده است لذا کم‌کم پای این ابزار جدید به علوم مهندسی آب و منابع طبیعی که جزء پدیده‌های غیرخطی و پیچیده هستند نیز باز شده است (۹).

با توجه به اینکه الگوریتم‌ها تنبل و مدل درخت تصمیم M5 سیستم‌های کارایی هستند و تاکنون تحقیقات کمتری در بررسی تراز سطح ایستابی با آن‌ها صورت گرفته است، در این مطالعه عملکرد دو مدل درخت تصمیم M5 و الگوریتم‌های تنبل KNN، LWL و KSTAR در پیش‌بینی تراز سطح ایستابی آبخوان تحت سناریوهای مختلف هواشناسی بررسی گردید.

مواد و روش:

شناخت منطقه مورد مطالعه: دشت نیشابور در زمینه آب زیرزمینی، یکی از دشت‌های مهم استان خراسان بوده و بعد از مشهد پرآب‌ترین دشت استان است، به طوری که سالانه بیش از ۷۰۰ میلیون مترمکعب آب زیرزمینی توسط چاه‌های عمیق و نیمه عمیق از آبخوان استخراج می‌شود (۱۱). از نظر موقعیت جغرافیایی حوزه مذکور در **حداصل ۱۳، ۵۸ تا ۳۰، ۵۹ طول شرقی و ۳۰، ۳۵ تا ۳۹، ۳۶ عرض شمالی** قرار دارد (۱۴). **حوزهی آبریز دشت نیشابور**، جزئی از حوضه‌ی آبریز کالشور نیشابور است که از شمال به **خط‌الرأس ارتفاعات بینالود**، از شرق به بلندی‌های لیلاجوق و یال پلنگ و از جنوب به **تپه‌ماهورهای نیزه بند**، سیاه کوه و کوه نمک و از غرب به **حوزهی آبریز دشت سبزوار محدود می‌شود** (۱۶).



شکل ۱. حوضه نیشابور و نمایی از رودخانه‌ها، چاه‌های پیژوهشی حوضه و آبخوان دشت نیشابور [۱۰]

روش KSTAR: یک روش طبقه‌بندی مبتنی بر مثال است که بر اساس موارد آموزشی مشابه، طبقه‌بندی را انجام می‌دهد و در مقایسه با الگوریتم‌های ماشین یادگیری، نتایج مطلوبی را حاصل می‌نماید. این روش برخلاف دیگر روش‌های داده کاوی که بر اساس تابع فاصله مبنی بر آنتروپی، طبقه‌بندی را انجام می‌دهند، از تابع شباهت برای تخمین متغیرهای مختلف استفاده می‌کند. فرض اصلی طبقه‌بندی مبتنی بر مثال، این است که موارد مشابه کلاس‌های مشابه داشته باشد (۲۱).

روش LWL: یک الگوریتم کلی برای یادگیری وزن‌دار شده به صورت محلی است. این الگوریتم با استفاده از روشی وزن‌دار را به نمونه‌ها نسبت می‌دهد و از روی نمونه‌های وزن‌دار شده، رده‌بندی را می‌سازد. این مدل برای مسائل طبقه‌بندی، رده‌بند Nave Bayes و برای مسائل رگرسیون، رگرسیون خطی انتخاب‌های خوبی هستند. می‌توان در این الگوریتم، تعداد همسایه‌های مورد استفاده را که پهنای باند هسته و شکل هسته مورد استفاده برای وزن‌دار کردن را (خطی، معکوس، یا گوسی) مشخص می‌کند، تعیین نمود (۱۰).

روش KNN: برخلاف توابع انتقالی کلاسیک، تکنیک k-نزدیک‌ترین همسایه از هیچ تابع ریاضیاتی از پیش تعریف شده‌ای

جهت تخمین متغیرهای مختلف استفاده نمی‌نماید (۳). در این روش، ابتدا نزدیک‌ترین همسایه‌های متغیر در ماه قبلی بر اساس فاصله اقلیدسی تعیین شده و همسایه‌ها بر اساس یک تابع وزن که بیشترین وزن را به نزدیک‌ترین همسایه می‌دهد باز نمونه‌گیری می‌شوند. این روش در واقع تابع توزیع احتمال شرطی ماه فعلی را نسبت به ماه قبلی تخمین می‌زند و شبیه‌سازی را با استفاده از آن انجام می‌دهد. به علاوه این روش می‌تواند اطلاعات اضافی ویژگی‌های بزرگ مقیاس اقلیمی را نیز به دست دهد. همچنین در شبیه‌سازی چند متغیر متغیرهای هواشناسی روزانه به کار رفت (۲۵).

روش درخت تصمیم M5: مدل M5 یک مدل درختی برای پیش‌بینی صفات عددی پیوسته است که در آن توابع رگرسیونی

خطی در برگ‌های این درخت متظاهر می‌شوند (۲۲). یک درخت تصمیم معمولاً از چهار بخش ریشه، شاخه، گره‌ها و برگ‌ها تشکیل شده است که گره‌ها با دایره نشان داده می‌شوند و شاخه‌ها نشان‌دهنده اتصال بین گره‌ها می‌باشند. درخت تصمیم به منظور سادگی در رسم معمولاً از چپ به راست و بالا از بالا به پایین کشیده می‌شود، به طوری که ریشه (گره اول) در بالا قرار می‌گیرد. انتهای یک زنجیره را برگ می‌نامند. هر گره مربوط به یک خصوصیت معین است و شاخه‌ها به معنای بازه‌ای از مقادیر هستند، این بازه‌های مقادیر بخش‌های مختلف مجموعه مقادیر معلوم را برای خصوصیت‌ها بدست دهند (۱۸). فرایند انشعاب در هر گره بارها تکرار می‌شود تا به گره پایانی یا همان برگ برسد که در برگ مجموع مجذور انحراف از میانگین داده‌ها حدوداً به صفر می‌رسد، با این کار درخت بزرگی توسعه پیدا خواهد کرد (۱۵).

جمع‌آوری داده‌ها و تجزیه و تحلیل آن‌ها: با استفاده از اطلاعات ۵۷ عدد چاه مشاهده‌ای از سال ۱۳۷۰ تا ۱۳۸۸ تراز سطح

آب دشت مورد بررسی قرار گرفت. در این مطالعه با توجه به آماربرداری چاه‌ها در سال‌های ۱۳۸۲ و ۱۳۸۸ از داده‌های تراز سطح ایستایی، بارندگی، تراز سطح زمین و تخلیه از سفره، برای پیش‌بینی تراز سطح ایستایی آبخوان استفاده شده است. ابتدا نرمال بودن داده با استفاده از نرم‌افزار GS⁺ و آرسی شد و بهترین مدل برازش داده‌ها از این نرم‌افزار تعیین و با استفاده از نرم‌افزار Arc GIS9.3 پهنه‌بندی صورت گرفت این کار برای دشت در پیکسل‌های ۴۵۰*۴۵۰ متر مربع صورت گرفت. مقدار تراز سطح ایستایی و بارندگی برای هر ماه در هر چاه استحصالی، با استفاده از نرم‌افزار Arc GIS9.3 تعیین شد (۵).

با توجه به پارامترهای تأثیرگذار بر سطح آب مانند داده‌های تغذیه و تخلیه سفره، تأخیر زمانی پارامترهای فوق و موقعیت دشت، داده‌های مورد استفاده در این مطالعه انتخاب شدند. این پارامترها عبارت‌اند از T تراز سطح زمین (متر)، Q حجم آب تخلیه‌شده استحصالی (هزار مترمکعب)، P t-1 مجموع بارندگی در ماه قبل (میلی‌متر)، P t-2 مجموع بارندگی در ۲ ماه قبل

(میلی متر)، p_t مجموع بارندگی در ماه (میلی متر). برای استفاده از داده‌های مذکور در الگوریتم‌های تنبل KSTAR, KNN و LWL

و درخت تصمیم M5، ابتدا داده‌ها به دودسته تقسیم شدند. ۷۵ درصد داده‌ها برای آموزش و ۲۵ درصد باقیمانده برای آزمون بکار

گرفته شد. خصوصیات آماری پارامترهای مورد استفاده در این تحقیق در جدول (۱) آمده است.

جدول ۱. خصوصیات آماری پارامترهای مورد استفاده

ماه	پارامتر	کمینه	بیشینه	میانگین	انحراف از معیار	ضریب تغییرات
مهر	تخلیه (هزار متر مکعب)	۱۲۱/۳۶	۶۷۶/۵۱۲	۰/۰۰۱۴	۱۳۴/۲۵	۱/۱۰۵
	بارندگی (میلی متر در ماه)	۱/۰۷	۱۶/۰۲	۰	۲/۱۳	۰/۰۱۹
	تراز سطح آب (متر)	۱۱۵۲/۸	۱۲۶۹/۱۲	۱۰۴۸	۳۲/۴	۰/۰۲۸۱
آبان	تخلیه (هزار متر مکعب)	۱۲۱/۳۶	۶۷۶/۵۱۲	۰/۰۰۱۴	۱۳۴/۲۵	۱/۱۰۵
	بارندگی (میلی متر در ماه)	۱۴/۱۴	۳۲/۹	۰	۸/۱۳	۰/۵۷
	تراز سطح آب (متر)	۱۱۵۲/۹۳	۱۲۶۸/۷	۱۰۴۸/۷۷	۳۲/۳	۰/۰۲۸
آذر	تخلیه (هزار متر مکعب)	۱۲۱/۳۶	۶۷۶/۵۱۲	۰/۰۰۱۴	۱۳۴/۲۵	۱/۱۰۵
	بارندگی (میلی متر در ماه)	۲۱/۶۶	۵۱/۱۴	۰/۷۹	۱۴/۶۷	۰/۶۷
	تراز سطح آب (متر)	۱۱۵۳/۳	۱۲۶۸/۷	۱۰۴۸/۶۷	۳۲/۴	۰/۰۲۸۱
دی	تخلیه (هزار متر مکعب)	۵۵/۲۵	۵۳۸/۲۸	۰/۰۰۷	۶۹/۳۴	۱/۲۵
	بارندگی (میلی متر در ماه)	۳۰/۰۱	۱۶۹/۳۵	۲/۲۴	۱۵/۶۹	۰/۵۲
	تراز سطح آب (متر)	۱۱۵۶/۲	۱۲۵۹/۵	۱۰۴۹/۶	۲۹/۶۶	۰/۰۲۵
بهمن	تخلیه (هزار متر مکعب)	۵۵/۲۵	۵۳۸/۲۸	۰/۰۰۷	۶۹/۳۴	۱/۲۵
	بارندگی (میلی متر در ماه)	۴۵/۶	۱۲۷/۸	۱/۰۹۲	۲۶/۰۶	۰/۵۷
	تراز سطح آب (متر)	۱۱۵۶/۴۳	۱۲۵۹/۵	۱۰۴۹/۷۷	۲۹/۶۶	۰/۰۲۵
اسفند	تخلیه (هزار متر مکعب)	۵۵/۲۵	۵۳۸/۲۸	۰/۰۰۷	۶۹/۳۴	۱/۲۵
	بارندگی (میلی متر در ماه)	۴۵/۷	۱۲۷/۸	۱/۰۹۲	۲۶/۰۷	۰/۵۷
	تراز سطح آب (متر)	۱۱۵۶/۲	۱۲۵۹/۵	۱۰۴۹/۶	۲۹/۴۶	۰/۰۲۵
فروردین	تخلیه (هزار متر مکعب)	۱۱۵/۴۸	۶۹۹	۰/۰۰۷۲	۱۳۱/۸	۱/۱۴
	بارندگی (میلی متر در ماه)	۷۴/۳۸	۱۶۷/۷	۱۵/۹۸	۲۸/۴۴	۰/۳۸
	تراز سطح آب (متر)	۱۱۵۶/۱	۱۲۵۹/۴۹	۱۰۴۹/۷۷	۲۹/۳۲	۰/۰۲۷
اردیبهشت	تخلیه (هزار متر مکعب)	۱۱۵/۴۸	۶۹۹	۰/۰۰۷۲	۱۳۱/۸	۱/۱۴
	بارندگی (میلی متر در ماه)	۴۵/۵	۱۱۲/۰۵	۱۹/۳۵	۱۹/۵	۰/۴۲
	تراز سطح آب (متر)	۱۱۵۶/۰۸	۱۲۵۹/۴	۱۰۴۹/۵	۲۹/۳	۰/۰۲۵
خرداد	تخلیه (هزار متر مکعب)	۱۱۵/۴۸	۶۹۹	۰/۰۰۷۲	۱۳۱/۸	۱/۱۴
	بارندگی (میلی متر در ماه)	۱۶/۹	۵۸/۸	۰/۱	۱۱/۱۶	۰/۶۵
	تراز سطح آب (متر)	۱۱۵۳/۵۸	۱۲۶۷/۷	۱۰۴۸/۸	۳۲/۳	۰/۰۲

۱/۱	۱۳۹/۵۶	۰/۰۰۷	۶۹۹	۱۲۶/۶۷	تخلیه (هزار مترمکعب)
۱/۹۳	۰/۴۶	۰	۲/۶۲	۰/۲۴	بارندگی (میلی متر در ماه)
۰/۰۲	۳۲/۳	۱۰۴۸/۹	۱۲۶۷/۶	۱۱۵۳/۴	تراز سطح آب (متر)
۱/۱	۱۳۹/۵۶	۰/۰۰۷	۶۹۹	۱۲۶/۶۷	تخلیه (هزار مترمکعب)
۰/۰۱	۲/۰۷	۰	۱۰/۱۲	۱/۴۹	بارندگی (میلی متر در ماه)
۰/۰۲۸	۳۲/۳۲	۱۰۴۸/۸	۱۲۶۷/۶	۱۱۵۳/۱۱	تراز سطح آب (متر)
۱/۱	۱۳۹/۵۶	۰/۰۰۷	۶۹۹	۱۲۶/۶۷	تخلیه (هزار مترمکعب)
۲/۵	۳/۸۹	۰	۳۱/۴۵	۱/۵	بارندگی (میلی متر در ماه)
۰/۰۲۸	۳۲/۳۲	۱۰۴۸/۴۶	۱۲۶۷/۶	۱۱۵۳/۰۳۹	تراز سطح آب (متر)
۰/۱	۱۳۳/۸	۱۰۶۸	۱۶۸۴/۵	۱۲۵۵/۴	تراز سطح زمین (متر)

قبل از استفاده از داده‌ها به منظور افزایش کارایی مدل‌ها، همه‌ی داده‌ها به شکل نرمال بین دو عدد ۰/۱ و ۰/۹ طبق فرمول (۱) استاندارد شدند (۲۳) آن‌ها بعد از استفاده در شبیه‌سازی به مقادیر اولیه برگشتند.

$$x_i = 0.8 \left(\frac{x - x_{min}}{x_{max} - x_{min}} \right) + 0.1 \quad -1$$

در این رابطه x_i مقدار استاندارد شده، x مقدار واقعی و x_{max} و x_{min} به ترتیب مقادیر حداقل و حداکثر داده‌ها می‌باشند. به منظور مقایسه نتایج حاصل از الگوریتم‌های تنبیل KNN، KSTAR، LWL و مدل درختی M5 در پیش‌بینی میزان تراز سطح ایستابی، هفت سناریو با توجه به پارامترهای تأثیرگذار بر سطح آب تعریف شد. بهترین سناریو موردنظر برای پیش‌بینی تراز سطح آب برای تمامی ماه‌های سال با استفاده از روش آزمون و خطا انجام شد. در جدول (۲) پارامترهای ورودی در هر سناریو ارائه گردید. برای بهره‌گیری از الگوریتم KNN، KSTAR، LWL و مدل درختی M5 از نرم‌افزار Weka 3.7 استفاده شده است.

جدول ۲. پارامترهای ورودی در سناریوهای مختلف

سناریو	پارامترهای ورودی
a	Q, P_t, P_{t-1}, T
b	Q, P_{t-1}, P_{t-2}, T
c	Q, P_{t-1}, T
d	Q, P_t, T
e	Q, T
f	Q, P_t
g	Q, P_t, P_{t-2}, T

T: تراز سطح زمین (متر)، Q: حجم آب تخلیه شده استحصالی (هزار مترمکعب)، Pt-1: مجموع بارندگی در ماه قبل

(میلی متر) Pt-2: مجموع بارندگی در دو ماه قبل (میلی متر) Pt: مجموع بارندگی در ماه (میلی متر)

معیارهای ارزیابی مدل: عملکرد معادلات تجربی، الگوریتم KNN و درخت تصمیم M5 توسط آماره‌های ریشه متوسط خطای مربعات (RMSE)، ضریب همبستگی (R^2) و متوسط قدر مطلق خطا (MAE) ارزیابی و از طریق رتبه‌بندی آماره‌ها و نزولی کردن آن‌ها در نرم‌افزار اکسل، بهترین رتبه انتخاب گردید.

$$RMSE = \sqrt{\frac{(\sum_{i=1}^n (E_{oi} - E_{pi})^2)}{n-1}} \quad -2$$

$$R^2 = \frac{(\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}))^2}{(\sum_{i=1}^n (x_i - \bar{x})^2 (\sum_{i=1}^n (y_i - \bar{y})^2)} \quad -3$$

$$MAE = \frac{1}{N} \sum_{i=1}^n (E_{oi} - E_{pi}) \quad -4$$

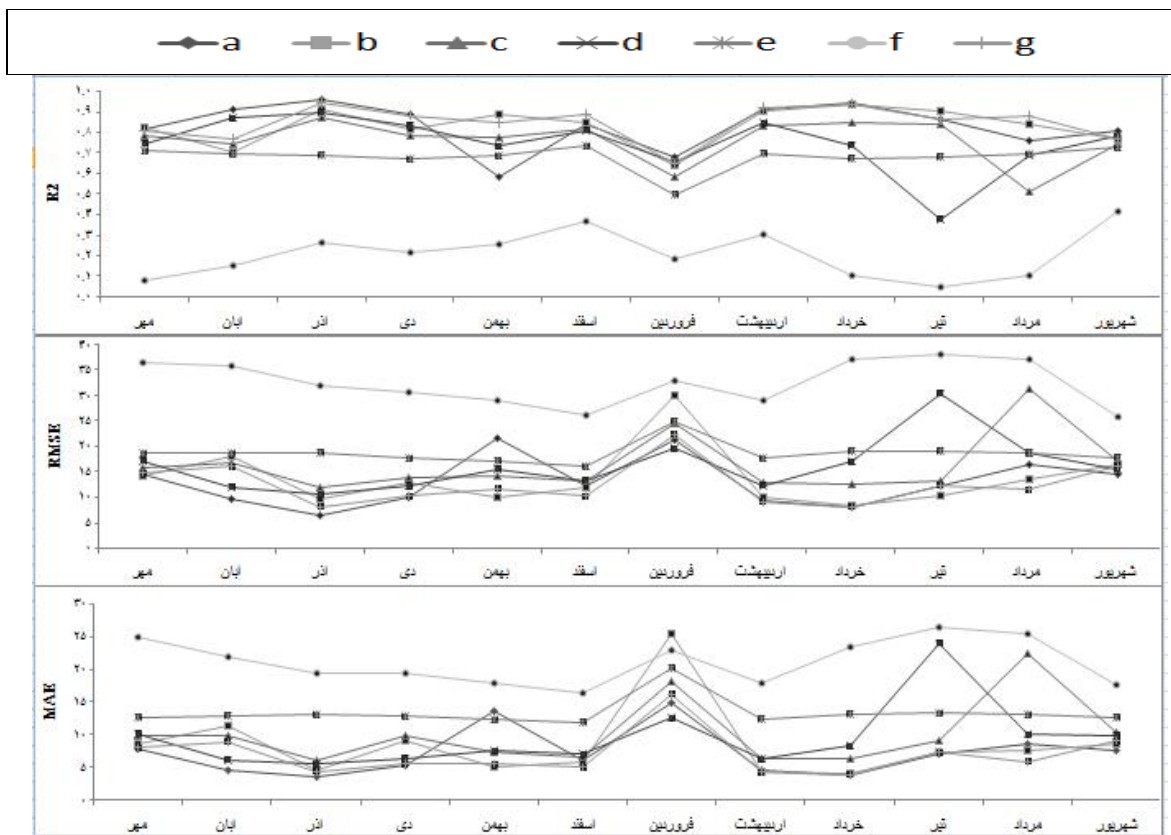
N تعداد کل داده‌ها، E_{oi} سطح آب تخمین زده شده، E_{pi} داده مشاهده‌ای با \bar{x} و \bar{y} متوسط مقادیر x_i و y_i هستند.

نتایج و بحث:

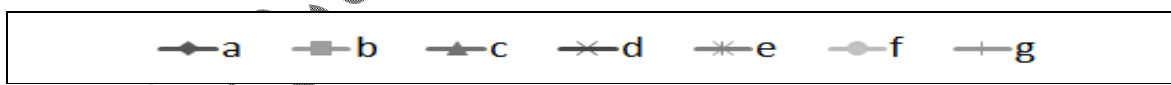
الگوریتم‌های تنبیل KNN، KSTAR و LWL و درخت تصمیم M5 تحت ۷ سناریو مورد بحث، در نرم‌افزار Weka اجرا گردید و شاخص‌های آماری به دست آمده از چهار مدل مذکور در شکل (۲) تا (۵) نمایش داده شد. همان‌طور که در شکل‌ها مشاهده می‌شود در تمام مدل‌ها، سناریو f در همه ماه‌ها دارای ضریب همبستگی کمتر و خطای بیشتری است که این نشان از کارایی کم این سناریو تحت تأثیر پارامترهای حجم آب تخلیه شده استحصالی و مجموع بارندگی در ماه، نسبت به سایر سناریوها است. علت این امر در نظر نگرفتن پارامتر تراز سطح زمین در این سناریو است. نتایج نشان می‌دهد که با تغییرات ارتفاع، میزان دبی چاه‌ها نیز تغییر می‌کند به طوری که دبی چاه‌ها در شمال و شرق دشت کمتر از جنوب و غرب دشت است و تغییرات خطوط تراز سطح آب نیز از خطوط تراز سطح زمین تبعیت می‌کند. این عامل می‌تواند نقش بسزایی در تخمین تراز سطح آب داشته باشد (۵). همچنین مشاهده می‌شود که در سناریوی a، b و g با اضافه شدن پارامترهای مجموع بارندگی در ماه قبل، مجموع بارندگی در دو ماه قبل و تراز سطح زمین، برآورد مطلوب تری از تراز سطح ایستابی حاصل می‌شود. این نتیجه با نتایج سیوکی و همکاران (۵) که به ارزیابی روش‌های آماری کلاسیک، سیستم‌های خبره (شبکه عصبی مصنوعی ANN و سیستم‌های استنتاج فازی عصبی (ANFIS) در پیش‌بینی سطح آب زیرزمینی دشت نیشابور پرداختند نیز مطابقت دارد. نتایج مدل‌های KNN، KSTAR و M5 بسیار به هم نزدیک می‌باشد اما مدل LWL در پیش‌بینی تراز سطح ایستابی ضعیف‌تر عمل می‌کند و در اکثر مواقع با تغییر سناریو تغییر محسوسی در نتایج این مدل حاصل نمی‌شود، بنابراین این مدل در پیش‌بینی تراز سطح ایستابی تحت سناریوهای مذکور مؤثر نخواهد بود.

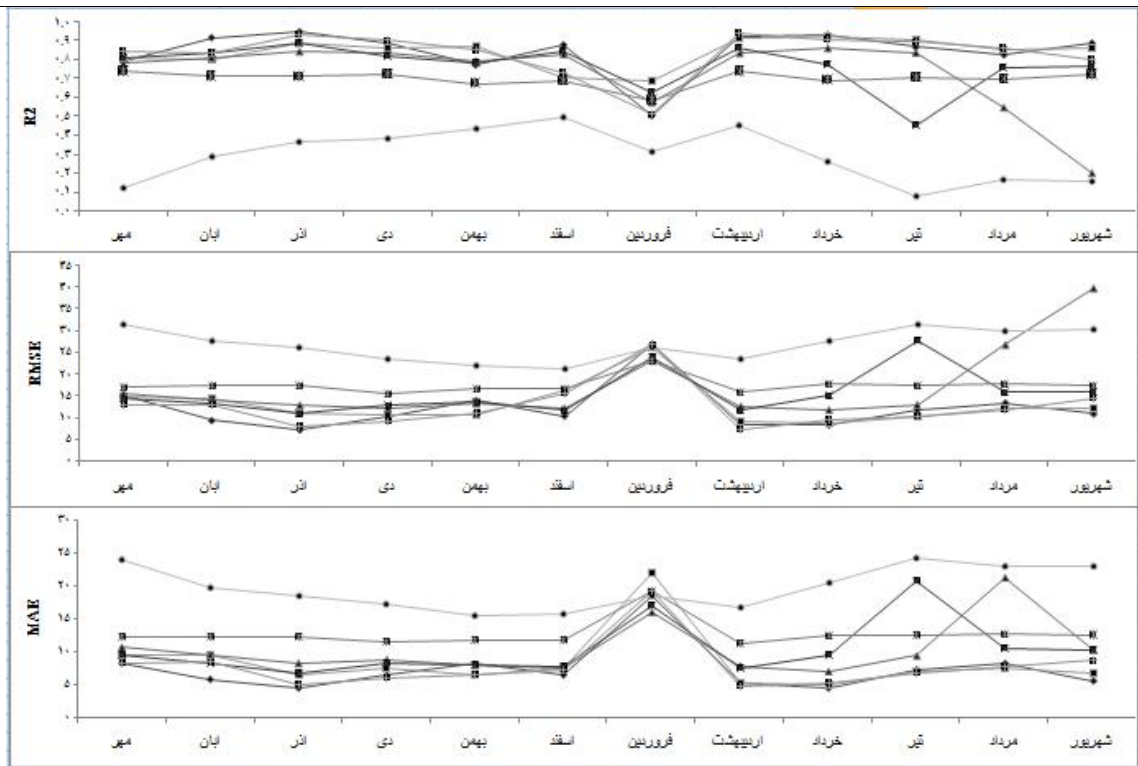
همچنین در تمامی مدل‌ها برآورد تراز سطح ایستابی در ماه فروردین دارای همبستگی کمتر و خطای بیشتری است که این به دلیل کم بودن داده‌های موجود در این ماه و حساسیت نرم‌افزار Weka به تعداد داده‌ها می‌باشد، به گونه‌ای که هرچه تعداد داده‌ها بیشتر

باشد پیش‌بینی دقیق‌تری را نتیجه می‌دهد. در ماه‌های فصل تابستان سناریوی c, d و e دارای ضریب همبستگی کمتر و خطای بیشتری می‌باشند. دقت مدل‌ها در پیش‌بینی تراز سطح ایستابی در ماه آذر در بین ماه‌های فصل پاییز بیشتر می‌باشد. این نتایج در هر سه مدل KNN، KSTAR و M5 مشاهده می‌شود.

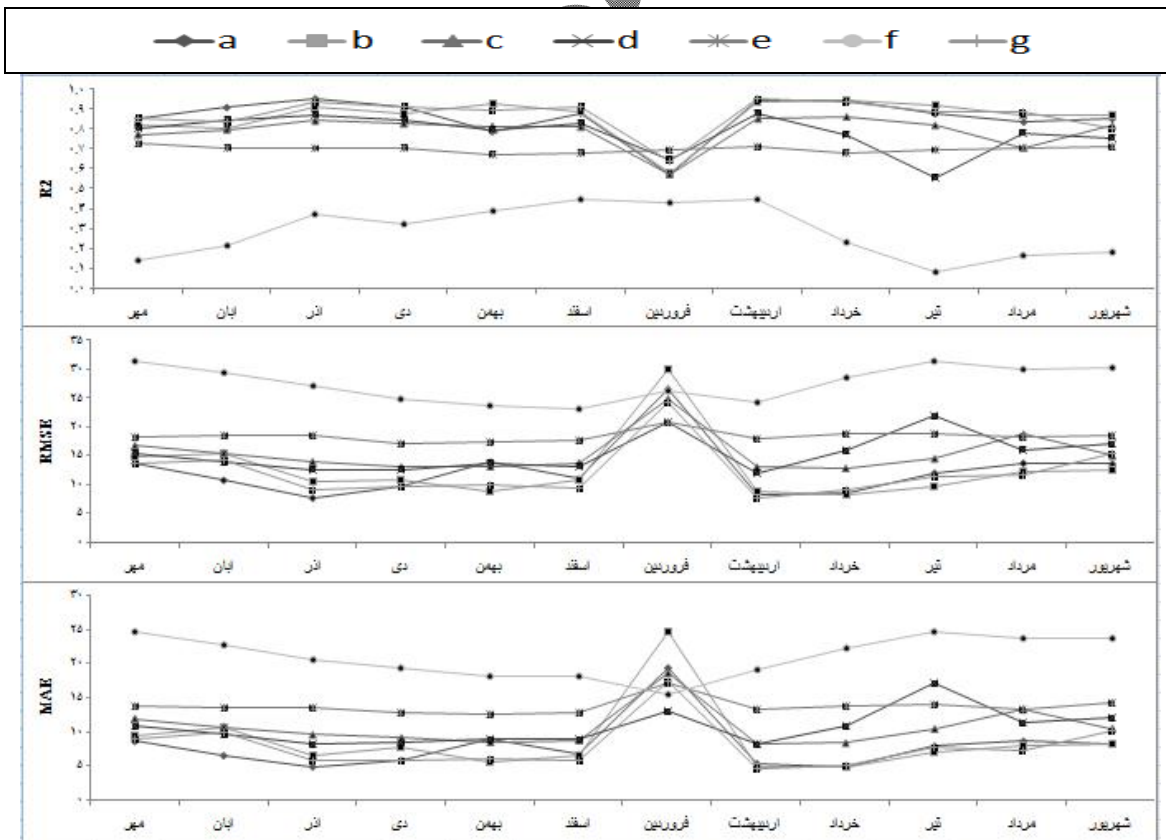


شکل ۲. نتایج مدل KNN در پیش‌بینی تراز سطح ایستابی آبخوان

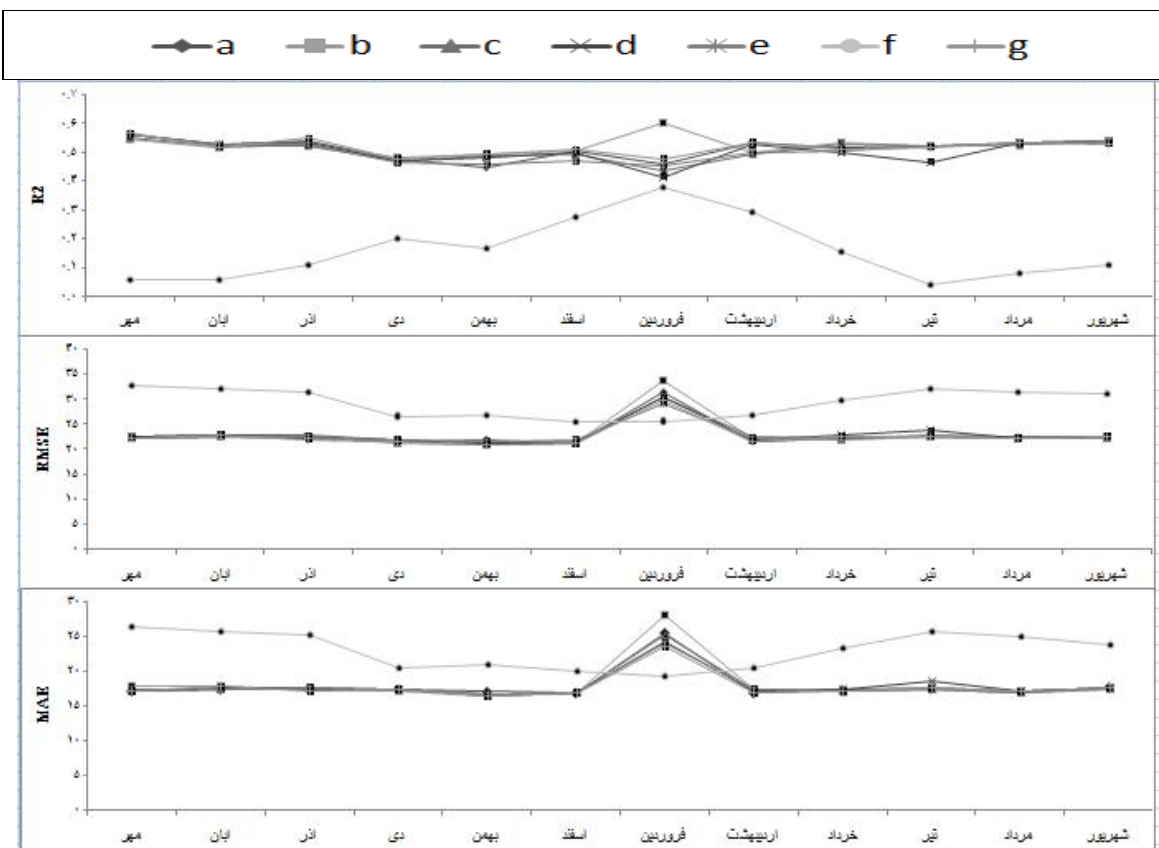




شکل ۳. نتایج مدل MS در پیش‌بینی تراز سطح ایستابی آبخوان



شکل ۴. نتایج مدل KSTAR در پیش‌بینی تراز سطح ایستابی آبخوان



شکل ۵. نتایج مدل LWL در پیش‌بینی تراز سطح ایستابی آبخوان

در جدول (۳) بهترین سناریو در هر مدل با توجه به بالا بودن ضریب همبستگی و کم بودن خطا و از طریق رتبه‌بندی آماره‌ها و نزولی کردن آن‌ها در نرم‌افزار اکسل، مشخص شده است. همان‌طور که مشاهده می‌شود در تمامی مدل‌ها به استثناء مدل LWL سناریو a در ماه آذر بهتر از ورودی‌های دیگر کارایی داشته است و در مدل LWL سناریو a در ماه بهمن نتایج مطلوب‌تری را نشان داده‌اند. با توجه به مطالب فوق می‌توان به این نتیجه رسید که زمانی که پارامترهای حجم آب تخلیه‌شده استحصالی، مجموع بارندگی در ماه، مجموع بارندگی در ماه قبل و تراز سطح زمین برای پیش‌بینی سطح آب زیرزمینی در نظر گرفته شود، نتایج دقیق‌تری حاصل می‌شود. همچنین در بین مدل‌های الگوریتم تنبل و مدل درخت تصمیم M5 کارایی مدل KNN تحت سناریوی a در ماه آذر با داشتن $R^2 = 0/96$ و $RMSE=6/56$ و $MAE=3/53$ از سایر مدل‌ها بیشتر است.

جدول ۳. تعیین بهترین سناریو برای هر مدل

مدل	سناریو	ماه	RMSE	R ²	MAE
KNN	a	آذر	۶/۵۶	۰/۹۶	۳/۵۳
KSTAR	a	آذر	۷/۵۴	۰/۹۵	۴/۸۸
LWL	b	بهمن	۲۰/۹۱	۰/۴۹	۱۶/۳۰

۱۶/۳۵	۰/۴۹	۲۰/۸۸	بهمن	g	
۴/۵۸	۰/۹۵	۷/۲۶	آذر	a	M5

در جدول (۴) بهترین مدل پیش‌بینی برای هر سناریو برحسب بالا بودن ضریب همبستگی و کم بودن خطا با ذکر نام ماه مشخص شده است و بهترین سناریوها به ترتیب رتبه مشخص گردیدند. همان‌طور که در قسمت قبل اشاره شد، سناریوی a در ماه آذر و با استفاده از مدل KNN به دلیل نظر گرفتن پارامترهای مجموع بارندگی در ماه قبل، مجموع بارندگی در دو ماه قبل و تراز سطح زمین، بیش‌بینی مطلوب‌تری را ارائه کرده است.

جدول ۴. تعیین بهترین مدل برای هر سناریو

رتبه	شماره سناریو	سناریو	ماه	مدل	RMSE	R^2	MAE
۱	۱	a	آذر	KNN	۶/۵۶	۰/۹۶	۳/۵۳
۳	۲	b	خرداد	KSTAR	۸/۲۷	۰/۹۴	۴/۹۰
۵	۳	c	آذر	KNN	۱۲/۰۴	۰/۸۷	۶/۱۹
۴	۴	d	آذر	KNN	۱۰/۶۹	۰/۹۰	۵/۶۱
۶	۵	e	اردیبهشت	M5	۱۵/۹۹	۰/۷۴	۱۱/۳۷
۷	۶	f	اسفند	M5	۲۱/۲۳	۰/۵۰	۱۵/۶۸
۲	۷	g	اردیبهشت	KSTAR	۷/۵۸	۰/۹۵	۴/۴۹

نتیجه‌گیری:

در این مطالعه توانایی الگوریتم‌های تنبلی KNN، KSTAR و LWL و مدل درخت تصمیم M5 در برآورد تراز سطح ایستایی آبخوان دشت بیرجند مورد ارزیابی قرار گرفت. نتایج به‌دست آمده حاکی از آن است که در تمام مدل‌ها، سناریو f در همه ماه‌ها به دلیل نظر نگرفتن پارامتر تراز سطح زمین، دارای ضریب همبستگی کمتر و خطای بیشتری می‌باشد و کارایی کمتری دارد. در سناریوی a، b و g با اضافه شدن پارامترهای مجموع بارندگی در ماه قبل، مجموع بارندگی در دو ماه قبل و تراز سطح زمین، برآورد مطلوب‌تری از تراز سطح ایستایی حاصل می‌شود. نتایج مدل‌های KNN، KSTAR و M5 بسیار به هم نزدیک می‌باشد اما مدل LWL در پیش‌بینی تراز سطح ایستایی ضعیف‌تر عمل می‌کند. زمانی که پارامترهای حجم آب تخلیه‌شده استحصالی، مجموع بارندگی در ماه، مجموع بارندگی در ماه قبل و تراز سطح زمین برای پیش‌بینی سطح آب زیرزمینی در نظر گرفته شود، نتایج دقیق‌تری حاصل می‌شود. همچنین در بین مدل‌های الگوریتم تنبلی و مدل درخت تصمیم M5 کارایی مدل KNN تحت سناریوی a در ماه آذر با $R^2 = 0.96$ ، $RMSE = 6.56$ و $MAE = 3.53$ از سایر مدل‌ها بیشتر می‌باشد.

منابع:

۱. ابارشی، ف.، م. مفتاح هلقی، ه. ثانی خانی و ا. ا. دهقانی. ۱۳۹۳. مقایسه عملکرد سه روش هوشمند به منظور پیش‌بینی نوسانات سطح ایستابی (مطالعه موردی: دشت زرین گل). مجله پژوهش‌های حفاظت آب و خاک. ۱۶۳(۱): ۱۸۰-۱۸۰.
۲. ایزدی، ع.، ا. داوری، ک.، ا. علیزاده، ب. قهرمان و س. ا. حقایقی مقدم. ۱۳۸۶. پیش‌بینی سطح ایستابی با استفاده از شبکه عصبی مصنوعی. مجله آبیاری و زهکشی ایران. ۲(۲): ۵۹-۷۱.
۳. جلالی، و.، م. همایی. ۱۳۹۰. ارائه مدلی غیر پارامتریک با استفاده از تکنیک k- نزدیک‌ترین همسایه در برآورد جرم مخصوص ظاهری خاک. مجله علوم و فنون کشاورزی و منابع طبیعی. علوم آب و خاک. ۱۵(۵۶): ۱۹۰-۱۸۱.
۴. حمزه، س.، ع. ع. ناصری و ح. ع. کشکولی. ۱۳۹۲. تغییرات پروفیل سطح ایستابی و دبی خروجی از زهکش‌های پلکانی در یک خاک مطبق. نشریه آب و خاک (علوم و صنایع کشاورزی). ۲۷(۱): ۱۳-۱.
۵. خاشعی سیوکی، ع.، ب. قهرمان و م. کوچک زاده. ۱۳۹۲. مقایسه مدل‌های شبکه عصبی مصنوعی، ANFIS و رگرسیون در برآورد سطح ایستابی آبخوان دشت نیشابور. نشریه آبیاری و زهکشی ایران. ۱(۷): ۲۲-۱۰.
۶. خلیلی پور، ا. ۱۳۸۱. بررسی کمی و کیفی آب‌های زیرزمینی دشت قم و تأثیر آن بر بیابان‌زایی منطقه. پایان‌نامه کارشناسی ارشد. دانشکده منابع طبیعی کرج. دانشگاه تهران. ۱۳۰ ص.
۷. زارع آبیانه، ح.، م. بیات ورکشی، ص. معروفی و ع. ایلدرومی. ۱۳۹۰. شبیه‌سازی سطح ایستابی دشت ملایر بر اساس داده‌های هواشناسی با استفاده از شبکه‌ی عصبی مصنوعی. پژوهش‌های جغرافیای طبیعی. ۷۸: ۲۸-۱۷.
۸. سلمی، ک. ۱۳۸۹. بررسی تغییرات سطح ایستابی در خاک با استفاده از مدل DRAINMOD و مدل شبکه عصبی مصنوعی در منطقه خوزستان. وب‌سایت سازمان هماهنگی مناطق استان تهران.
۹. طالبی، ع.، ز. اکبری. ۱۳۹۲. بررسی کارایی مدل درختان تصمیم‌گیری در برآورد رسوبات معلق رودخانه‌ای (مطالعه موردی: حوضه سد ایلام). مجله علوم و فنون کشاورزی و منابع طبیعی، علوم آب و خاک. ۱۷(۶۳): ۱۰۹-۱۲۱.
۱۰. عطایی، ش. ۱۳۸۶. نرم‌افزار داده‌کاوی Weka. دانشگاه علم و صنعت ایران، دانشکده مهندسی راه‌آهن.
۱۱. فرج زاده، م.، س. ولایتی و آ. حسینی. ۱۳۹۱. تحلیل بحران آب در دشت نیشابور با رویکرد برنامه‌ریزی محیطی. شرکت سهامی مدیریت منابع آب ایران، دانشگاه تربیت مدرس تهران.
۱۲. محتشم، م.، ا. ا. دهقانی، ا. اکبرپور، م. مفتاح هلقمی و ب. اعتباری. ۱۳۸۹. پیش‌بینی سطح ایستابی با استفاده از شبکه عصبی مصنوعی (مطالعه موردی: دشت بیرجند). مجله آبیاری و زهکشی ایران، ۱(۴): ۱-۱۰.
۱۳. مختاری، ز.، ا. ناظمی، ا. اصغری مقدم و ع. ندیری. ۱۳۹۰. پیش‌بینی زمانی و مکانی تراز آب زیرزمینی با استفاده از کریجینگ عصبی مطالعه موردی: دشت شبستر. سی‌امین گردهمایی علوم زمین.
۱۴. نیک‌نیا، ع. ۱۳۹۲. ارزیابی ریسک آلودگی سد بار با مدل وراست یک در جی‌آی‌اس. فصل سوم پایان‌نامه.
۱۵. نهرین، ف.، م. ت. ستاری و ن. بیگ زالی. ۱۳۹۲. مقایسه تخمین بار معلق به دو روش منحنی سنج رسوب و مدل درختی M5 (مطالعه موردی: رودخانه ليقوان چای). دوازدهمین اجلاس هیدرولیک ایران. گروه مهندسی آبیاری و آبادانی، پردیس کشاورزی و منابع طبیعی.

۱۶. ولایتی، س. ۱۳۷۹. مهم ترین عوامل مؤثر بر تغییرات کیفی آبخوان، دشت نیشابور. فصلنامه تحقیقات جغرافیایی، ۱۵(۳): ۱۱۹-۱۳۴.
17. Djurović, N., R. Stričević and R. Gajić. 2000. Some constraints of the application of methods for drain spacing determination in unsteady- state of flow in eugley soil. *Journal of Agricultural Sciences*. 45 (2): 83-91.
 18. Fallahi, M.R., H. Varvani and S. Goliyan. 2012. Precipitation forecasting using regression tree model to flood control. 5 th "International watershed and water and soil resources management, Kerman, Iran.
 19. Foroghi, H., H. Moazed., M. Kashefipour and M. Karami. 2006 . Management Level Maron River Basin Aystaby Using Artificial Neural Networks .Seventh Seminar on International River Engineering, Ahvaz.
 20. Hand, D., M. Heikki and S. Padhraic. 2001. Principles of Data Mining. The MIT press.
 21. Cleary .J.G and Trigg. L.E. 1995. K*: An Instance- based Learner Using an Entropic Distance Measure .Proceedings of the 12th International Conference on Machine learning, pp. 108-114.
 22. Pal. M. 2006 . M5 model tree for land cover classification. *International Journal of Remote Sensing*. 27(4):825-831.
 23. Rahimi Khoob, A. 2008. Artificial neural network estimation of reference evapotranspiration from pan evaporation in a semi-arid environment. *Journal of Irrigation Science*. 27(1): 35-39.
 24. Sabti, N.A. 1989. Linear and nonlinear solution of the Boussinesq equation for the bi-level drainage problem. *Agric. Water Manage*. 16:269-278.
 25. Yates, D., S. Gangopadhyay., B. Rajagopalan and K. Strzepek. 2003. A technique for generating regional climate scenarios using a nearest-neighbor algorithm. *water resources research*. 39 (7):1-15.

Compare lazy algorithms and M5 model to estimate Groundwater Level (Case Study of Plain Neyshabur)

Abstract:

In recent years and in many countries, overusing groundwater resources had been higher than their annual feeding amount. This issue caused lowering the groundwater levels, followed by drying wells, qanats and springs. In this study, given the importance of Neyshabur plain in supplying agricultural, industrial and drinkable water of the area, lazy algorithms of KNN, KSTAR and LWL and M5 tree model have been utilized under seven different scenarios in order to estimate groundwater level of this aquifer. To compare the results, the Statistical parameters of root mean square error, correlation coefficient and the average absolute error were analyzed. The results showed that the 'f' scenario which contains the volume of water discharged and total precipitation parameters, had lower efficiency because the ground surface level parameter was not taken into account. In 'a', 'b' and 'g' scenarios, an optimum estimation has been maintained for the groundwater level due to the considering of the parameters of total rainfall in the previous month, total rainfall in the last two months and ground surface level. Among the models of lazy algorithms and M5 decision tree model, the ability of KNN model under 'a' scenario was higher than other models in December (Azar) by the statistical parameters of $R^2 = 0.96$, RMSE= 6/56 and MAE= 3/53. Also check out evaluation criteria showed that the LWL is not an appropriate model to predict the level of the water table.

Keywords: lazy algorithm, M5 tree model, the static surface level, Neyshabur plain

با سلام و احترام

پاسخ به سوالات داوران:

متنی که ملاحظه می‌کنید در بر گیرنده پاسخ سوالات داوران در مورد مقاله "مقایسه الگوریتم تنبل و مدل M5 در برآورد تراز سطح ایستابی (مطالعه موردی دشت نیشابور)" می‌باشد.

نویسندگان لازم می‌دانند از تمامی داوران بخاطر وقتی که برای مطالعه و داوری مقاله صرف کردند تشکر کنند. به منظور افزایش کیفیت مقاله، نظرات داوران به صورت کامل در مقاله لحاظ شده است. به همین دلیل انتظار می‌رود بازنگری‌های صورت گرفته در مقاله باعث بهبود مقاله در تمامی سطوح شده باشد. تمامی اصلاحات بر روی فایل مقاله به صورت رنگی مشخص شده است.

در ادامه متن تمامی سوالات داوران پاسخ داده شده است.

مقاله پذیرفته شده، ویراستاری نشده