



## Evaluation of Support Vector Machine, and Linear and Quadratic Discriminant Analysis for Groundwater Level Variations in Shahrekord Plain

A. Ramezani-Charmahineh<sup>1\*</sup>  
and M. Zounemat-Kermani<sup>2</sup>

### Abstract

In recent years, due to the increasing rate of water demand and severe droughts, groundwater resources are considered as the most important sources of fresh water. Accordingly, a comprehensive strategy along with a long term plan is needed for preventing groundwater destruction. Variations in aquifer water level, are amongst the main factors which provide correct judgment about groundwater status and govern the watershed management projects. In the present study, monthly data (1999 to 2009) from 33 observational wells in Shahrekord Plain have been used for simulating the groundwater level. The relationship among the Shahrekord Plain coordinates and the groundwater level variations, for 1, 3, 5 and 10 year period, were investigated using Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), and Support Vector Machine (SVM). The results showed that the SVM is superior to the other two models due to its lowest average relative error in 1 and 3 year periods, and its acceptable precision in 5 and 10 year periods.

**Keywords:** Groundwater level, Linear discriminant analysis, Shahrekord Plain, Support vector machine.

Received: September 6, 2014

Accepted: February 19, 2015

## ارزیابی روش‌های ماشین بردار پشتیبان، تحلیل تفکیک خطی و درجه دوم در شبیه‌سازی نوسانات سطح آب زیرزمینی دشت شهرکرد

عبداله رضانی چرمهینه<sup>۱\*</sup> و محمد ذونعمت کرمانی<sup>۲</sup>

### چکیده

با توجه به افزایش تقاضای آب و خشکسالی‌های شدید در سال‌های اخیر، منابع آب زیرزمینی به عنوان مهم‌ترین منبع تأمین آب شیرین دارای اهمیت زیادی هستند. لزوم توجه به این منابع و جلوگیری از تخریب آنها نیازمند به رویکرد جامع و برنامه‌ریزی بلند مدت دارد. تغییرات آبخوان عامل مهمی است که امکان قضاوت صحیح در مورد روند تراز سطح آب زیرزمینی و مدیریت لازم در حوزه آبخیز را ایجاد می‌کند. در این تحقیق از اطلاعات ماهانه مربوط به ۳۳ چاه مشاهده‌ای در دشت شهرکرد طی سال‌های ۱۳۷۸ تا ۱۳۸۸ به منظور شبیه‌سازی سطح آب زیرزمینی استفاده شده است. ارتباط مختصات دشت شهرکرد و تغییرات سطح آب زیرزمینی با روش‌های تحلیل تفکیک خطی (LDA)، تحلیل تفکیک درجه دوم (QDA) و ماشین بردار پشتیبان (SVM) در دوره‌های زمانی ۱، ۳، ۵ و ۱۰ ساله مورد بررسی قرار گرفته است. نتایج نشان از این دارد که ماشین بردار پشتیبان با داشتن کمترین میانگین خطای نسبی در دوره‌های ۱ و ۳ ساله و دقت قابل قبول در دوره‌های ۵ و ۱۰ ساله دارای برتری بر سایر روش‌ها بوده است.

**کلمات کلیدی:** تراز آب زیرزمینی، تحلیل تفکیک خطی، دشت شهرکرد، ماشین بردار پشتیبان.

تاریخ دریافت مقاله: ۹۳/۶/۱۵

تاریخ پذیرش مقاله: ۹۳/۱۱/۳۰

1-M.Sc. Graduated, Department of Water Engineering, Shahid Bahonar University of Kerman, Kerman, Iran. Email: ramezaniwater@yahoo.com

2- Assistant Professor, Department of Water Engineering, Shahid Bahonar University of Kerman, Kerman, Iran.

\*- Corresponding Author

۱- دانش‌آموخته کارشناسی ارشد، بخش مهندسی آب، دانشگاه شهید باهنر کرمان، کرمان، ایران.

۲- استادیار، بخش مهندسی آب، دانشگاه شهید باهنر کرمان، کرمان، ایران.

\*- نویسنده مسئول

## ۱- مقدمه

آب زیرزمینی آبخوان ساحلی در بندر Mukho کره جنوبی بررسی کردند. نتایج حاکی از کارایی دو روش بود به طوری که مدل شبکه عصبی مصنوعی در مراحل آموزش و آزمون، جذر میانگین مربعات خطای کمتری داشت اما ضریب راندمان بیشتر در ماشین بردار پشتیبان موجب عملکرد بهتر آن گردید. (Sudheer et al. (2011) تراز سطح آب زیرزمینی منطقه Rentachintala در هندوستان را مورد مطالعه قرار دادند. ماشین بردار پشتیبان با به کارگیری الگوریتم بهینه‌سازی انبوه ذرات با رفتار کوانتومی در تنظیم مقادیر پارامترها، عملکرد بهتری در مقایسه با شبکه عصبی مصنوعی در پیش‌بینی تراز سطح آب زیرزمینی از خود نشان داد. (Zhao et al. (2012) کارایی روش‌های ماشین بردار پشتیبان و ماشین بردار ارتباط بر پایه الگوریتم بهینه‌سازی انبوه ذرات را در مدل‌سازی سطح آب زیرزمینی ماهانه Beijing در چین مورد بررسی قرار دادند. مدل ماشین بردار ارتباط از دقت بیشتری برخوردار بوده است. (Shiri et al. (2013) به مقایسه روش‌های برنامه‌ریزی بیان ژن، سیستم استنتاج تطبیقی عصبی-فازی، شبکه عصبی مصنوعی، ماشین بردار پشتیبان و خود همبسته-میانگین متحرک در پیش‌بینی نوسانات سطح آب زیرزمینی ایستگاه Hongcehon در کره جنوبی پرداختند. نتیجه مقایسه مدل‌ها، برتری مدل برنامه‌ریزی بیان ژن بر مدل‌های دیگر بود. (Suryanarayana et al. (2014) نوسانات ماهانه سطح آب زیرزمینی مناطق Sivajipalem, Madhurawada و Gullalapalem در هندوستان را با مدل‌های رگرسیون بردار پشتیبان، شبکه عصبی مصنوعی، خود همبسته-میانگین متحرک تلفیقی و مدل ترکیبی موجک-رگرسیون بردار پشتیبان مورد پژوهش قرار دادند که کارایی بیشتر مدل موجک-رگرسیون بردار پشتیبان نسبت به سایر مدل‌ها را به همراه داشت. توجه به اکثر مطالعات پیشین، بیانگر به کارگیری روش ماشین بردار پشتیبان در مقایسه با سایر روش‌های فراکاووشی (شبکه عصبی مصنوعی، عصبی موجک و غیره) می‌باشد. این در حالی است که تاکنون با توجه به اطلاعات نگارندگان ارزیابی عملکرد روش ماشین بردار پشتیبان (که در زمره روش‌های فراکاووشی قرار می‌گیرد) در قیاس با روش‌های تحلیل تفکیکی (که مبنای آماری دارند) در ارتباط با مدل‌سازی آب زیرزمینی مورد بررسی قرار نگرفته است. همچنین لازم به یادآوری است که آبخوان‌ها از اهمیت زیادی در منابع آب برخوردار هستند و تغییرات سطح آب زیرزمینی عامل مهمی در استفاده از آبخوان‌ها و مدیریت صحیح آن‌ها می‌باشد. بنابراین در این تحقیق استفاده از روش‌های یاد شده در شبیه‌سازی تغییرات سطح آب زیرزمینی مورد بررسی قرار گرفته است.

کشور ایران با توجه به اقلیم نامناسب و کمبود ریزش‌های جوی، از جمله کشورهای کم آب جهان محسوب می‌گردد به طوری که متوسط بارندگی سالانه آن حدود یک سوم جهان می‌باشد (باقری‌هارونی و فتحیان، ۱۳۹۰). از طرفی به موازات رشد سریع جمعیت و افزایش نیازهای بشری، تأمین آب برای مصارف شرب، کشاورزی و صنعت از اهمیت ویژه‌ای برخوردار گردیده است. در این میان آب‌های زیرزمینی به منزله منابعی مطمئن جهت تأمین آب تلقی می‌گردند (کوهستانی، ۱۳۹۰). شبیه‌سازی تغییرات سطح آب زیرزمینی از مباحث مهم در منابع آب می‌باشد که می‌توان با توجه به آن، برنامه‌ریزی‌های لازم در این زمینه را انجام داد. همچنین با وجود روند تشدید مشکلات مربوط به آب و افزایش رو به رشد تقاضای آن، مدیریت موفق منابع آب نیازمند رویکردهای جهت‌دار، جامع و هدفمند می‌باشد تا بتواند نیاز مصرف‌کنندگان را برطرف سازد. در سال‌های اخیر نیز رشد روزافزون کاربرد هوش محاسباتی در حل مسائلی دیده می‌شود که تأثیر پارامترهای بسیار زیاد در برخی فرایندهای فیزیکی و وجود روابط کاملاً غیرخطی میان آنها حاکم می‌باشد و مبنای روش‌های هوشمند، استفاده از دانش نهفته در داده‌ها، تلاش برای استخراج روابط ذاتی بین آنها و تعمیم آن در موقعیتهای دیگر است (محرّمپور و همکاران، ۱۳۹۰). بنابراین بهره‌گیری از شیوه‌های نوین شبیه‌سازی منابع آب از جمله روش ماشین بردار پشتیبان (SVM)<sup>۱</sup> دارای جایگاه مهمی می‌باشد. همچنین جهت مدیریت جامع‌تر آبخوان‌ها، توجه به روش‌های آماری هم از اهمیت ویژه‌ای برخوردار است. از جمله این روش‌ها، تحلیل تفکیکی خطی (LDA)<sup>۲</sup> و تحلیل تفکیکی درجه دوم (QDA)<sup>۳</sup> را می‌توان نام برد.

در ارتباط با ادبیات فنی موضوع مورد نظر، محققانی چند به مدل‌سازی و ارزیابی عملکرد روش ماشین بردار پشتیبان و تحلیل تفکیکی پرداخته‌اند. (Behzad et al. (2010) به مقایسه مدل‌های ماشین بردار پشتیبان و شبکه عصبی در شبیه‌سازی سطح آبخوان Towaco در آمریکا در دوره‌های روزانه، هفتگی، دو هفته، ماهانه و دو ماهه پرداختند. نتایج نشان از برتری ماشین بردار پشتیبان بر شبکه عصبی به خصوص در دوره‌های زمانی طولانی‌تر داشت. Rani et al. (2011) تغییرات فصلی در کیفیت آب را در سه انشعاب اصلی رود Ganga شامل رودخانه‌های Gandak, Ghaghra و Sone در جلگه Gangetic در هندوستان با استفاده از روش تحلیل تفکیکی مورد تحقیق قرار دادند. (Yoon et al. (2011) روش‌های ماشین بردار پشتیبان و شبکه عصبی مصنوعی را برای پیش‌بینی تراز سطح

۲- مواد و روش‌ها

۲-۱- منطقه مورد مطالعه

ایجاد می‌کند. مدل شامل عملکرد توصیفی بر پایه ترکیبات متغیرهای پیش‌بینی کننده است که توصیف قابل قبولی بین گروه‌ها ایجاد می‌کند. این روش در موارد قابل اندازه‌گیری و قابل پیش‌بینی که رابطه بین گروه‌ها ناشناخته است کاربرد دارد (Huberty, 1994). در این روش به صورت خودکار دو تابع انتخاب می‌شود. در تابع اول اعضای هر گروه که دارای تأثیر زیادی بر روی گروه‌های مطالعاتی دارند توسط مدل جدا می‌شوند. در تابع دوم، مدل اعضای انتخابی که دارای عدم تشابه با اعضای تابع اول و در صورت امکان تأثیر زیاد بر روی گروه‌های مطالعاتی هستند را انتخاب می‌کند (پورکرمان و همکاران، ۱۳۹۰).

۲-۲-۱-۱- تحلیل تفکیک خطی و تحلیل تفکیک درجه دوم

تحلیل تفکیک خطی و درجه دوم روش‌های آماری هستند که برای پیدا کردن ترکیب خطی و درجه دوم خصوصیتی که به بهترین صورت دو یا چند گروه از اشیا را از هم جدا می‌کنند، استفاده می‌شوند (چینی‌پرداز و همکاران، ۱۳۸۵). در مسأله تفکیک، مجموعه راهنما  $\{x_i, g_i\}_{i=1}^n$  و توزیع  $x_i$  به صورت  $X_i \sim N_p(\mu_j, \Sigma)$  برای  $j=1, 2, \dots, J$  در نظر گرفته می‌شوند. پس  $\varphi$  که چگالی شرطی  $x$  در گروه  $j$  ام می‌باشد از رابطه ۱ به دست می‌آید.

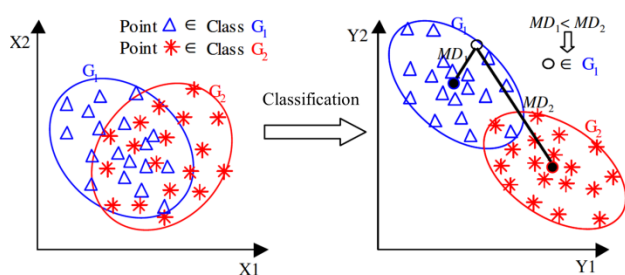
$$\varphi(\mu_j, \Sigma) = P(X = x | G = j) = \frac{1}{2\pi|\Sigma|} \exp\left\{-\frac{1}{2}MD(x_i, \mu_j)\right\} \quad (1)$$

طبق رابطه ۲ نیز لگاریتم درستنمایی تعریف می‌شود که  $\sum_{g_i=j}$ ، مجموع مشاهدات در گروه  $j$  ام را نشان می‌دهد.

$$2\log(\mu_j, \Sigma) = -\sum_{j=1}^J \sum_{g_i=j} MD(x_i, \mu_j) - n \log |\Sigma| \quad (2)$$

$$MD(x_i, \mu_j) = (x_i - \mu_j)' \Sigma^{-1} (x_i - \mu_j) \quad (3)$$

رابطه ۳، فاصله Mahalanobis بین  $x$  و  $\mu$  را مشخص می‌کند که با استفاده از آن جداسازی گروه‌ها (کلاس) انجام می‌گیرد (شکل ۲).



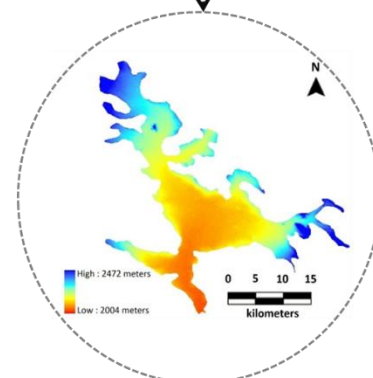
شکل ۲- جداسازی گروه‌ها بر اساس فاصله Mahalanobis (Du and Jin, 2008)

دشت شهرکرد با وسعتی نزدیک به ۵۵۱ کیلومترمربع در ناحیه زاگرس بلند و استان چهارمحال و بختیاری و نیز در طول جغرافیایی  $33^{\circ} 35'$  تا  $32^{\circ} 07'$  شرقی و عرض جغرافیایی  $51^{\circ} 10'$  تا  $50^{\circ} 38'$  شمالی قرار گرفته‌است. این دشت دارای ۴۱۷ حلقه چاه کشاورزی، ۵۹ حلقه چاه آب شرب، ۱۵۹ حلقه چاه صنعتی، ۷۹ رشته قنات و ۴۰ دهنه چشمه فعال است که سالانه حدود ۲۳۰ میلیون مترمکعب از منابع آب زیرزمینی دشت را تخلیه کرده و به مصارف گوناگون می‌رساند. درصد بسیار بالایی از آن (بیش از ۹۰ درصد) در بخش کشاورزی در فصل زراعی استفاده می‌شود. بخشی از آب شرب مردم شهرستان شهرکرد نیز از همین چاه‌ها تأمین می‌شود (لاله‌زاری و طباطبایی، ۱۳۸۹). موقعیت جغرافیایی و نقشه ارتفاعی دشت شهرکرد در شکل ۱ نشان داده شده است. وضعیت توپوگرافی نشان داده شده در شکل ۱، بیانگر مرتفع بودن نواحی شمال غربی و شرقی است. در حالی که نواحی مرکزی و جنوبی دشت دارای کمترین ارتفاع می‌باشند.

الف



ب

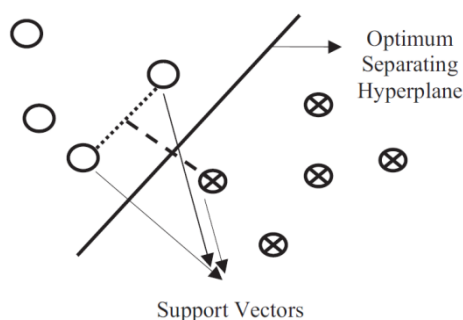


شکل ۱- دشت شهرکرد؛ (الف): موقعیت جغرافیایی، (ب): نقشه ارتفاعی

۲-۲- روش‌های مورد استفاده در این تحقیق

۲-۲-۱- تحلیل تفکیکی

تحلیل تفکیکی یک مدل پیش‌بینی کننده برای اعضای یک گروه



شکل ۳- جداسازی گروه‌ها با بیشینه‌سازی فاصله بین داده‌های دو گروه توسط یافتن ابرصفحه جداکننده بهینه (Güraksm et al., 2014)

ضرایب لاگرانژ  $(\lambda_i^- \text{ و } \lambda_i^+)$  به دست می‌آیند. برای تعدادی از  $x_i$  ها شرط  $(\lambda_i^- - \lambda_i^+) \neq 0$  برقرار می‌باشد که بیانگر بردارهای پشتیبان می‌باشند. تابع پیش‌بینی با استفاده از بردارهای پشتیبان ساخته می‌شود که در رابطه ۸ مشخص شده است و  $k(x_i, x)$  تابع کرنل می‌باشد.

$$f(x) = \sum_{i=1}^N (\lambda_i^- - \lambda_i^+) k(x_i, x) + b \quad (8)$$

تابع پایه شعاعی (RBF) از توابع کرنل شناخته شده در SVM می‌باشد (Chen and Yu, 2007) که در رابطه ۹ مشخص شده است.

$$k(u, v) = \exp\left(-\frac{\|u - v\|^2}{2\sigma^2}\right), \quad \frac{1}{2\sigma^2} = \gamma \quad (9)$$

### ۳-۲- بررسی ارتباط مختصات دشت شهرکرد با نوسان سطح آب زیرزمینی

دشت شهرکرد به دلیل میزان بارش و برداشت آب متفاوت در نواحی مختلف با نوسان سطح آب مواجه شده است. شبیه‌سازی کاهش یا افزایش سطح آب زیرزمینی می‌تواند کمک شایانی در تحلیل روند تغییرات سطح آب دشت نماید. بنابراین مختصات چاه‌های مشاهده‌ای دشت به عنوان ورودی برای مدل‌های LDA، QDA و SVM نظر گرفته شده است و بر اساس این ورودی گروه‌های (کلاس) تعریف شده برای میزان تغییرات سطح آب، تعیین گشته است. برای شبیه‌سازی نوسان سطح آبخوان، اطلاعات سطح آب زیرزمینی در طی سال‌های ۱۳۷۸ تا ۱۳۸۸ مورد استفاده قرار گرفته است. دوره‌های زمانی ۱ ساله (۱۳۸۷ تا ۱۳۸۸)، کوتاه مدت ۳ سال (۱۳۸۵ تا ۱۳۸۸)، میان مدت ۵ سال (۱۳۸۳ تا ۱۳۸۸) و بلند مدت ۱۰ سال (۱۳۷۸ تا ۱۳۸۸) برای بررسی روند نوسان انتخاب شده‌اند. موقعیت

احتمال پیشین گروه  $j$  ام به صورت  $P(G=j) = \prod_j$  در نظر گرفته می‌شود که معمولاً از قبل مشخص است یا از مجموعه راهنما برآورد می‌شود. در صورتی که هیچ دلیل قبلی بر برتری گروه‌ها بر یکدیگر نباشد  $P(G=j) = \frac{1}{J}$  فرض می‌شود. حالت غیرخطی و تابع درجه دوم از  $x$  به صورت رابطه ۴ می‌باشد.

$$x'(\sum_i^{-1} \sum_j^{-1})x \quad (4)$$

### ۲-۲-۲- ماشین بردار پشتیبان

SVM از روش‌های یادگیری با نظارت است که از آن برای گروه‌بندی و برآورد تابع برازش داده‌ها استفاده می‌شود به طوری که کمترین خطا در گروه‌بندی داده‌ها یا تابع برازش رخ دهد. این روش بر اساس تئوری یادگیری آماری<sup>۴</sup> است و از اصل کمینه‌سازی خطای ساختاری (SRM)<sup>۵</sup> بهره می‌گیرد تا یک جواب بهینه کلی حاصل شود (Vapnik, 1998).

در ابتدا مجموعه آموزشی  $m$  الگویی با مقادیر پیش‌بینی طبق رابطه ۵ مقایسه می‌شود.

$$T = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\} \quad x \in R^m, \quad y \in R \quad (5)$$

که  $x_i$  پارامتر یا ویژگی ورودی،  $y_i$  مقدار هدف و  $m$  تعداد ویژگی‌های ورودی است. در مدل SVM از مجموعه توابع (ابرفضحه) به فرم  $f(x) = wx + b$  برای پیش‌بینی استفاده می‌شود که  $w$  وزن بردار  $x$  و  $b$  مقدار bias می‌باشند. طبق رابطه ۶ برای کمینه کردن خطای آزمون باید عبارت پیچیدگی حداقل شود که مستلزم کمینه‌سازی norm بردار وزن است (رابطه ۷). این کمینه‌سازی معادل با بیشینه‌سازی فاصله بین داده‌های دو گروه توسط یافتن ابرصفحه جداکننده بهینه<sup>۶</sup> می‌باشد (شکل ۳).

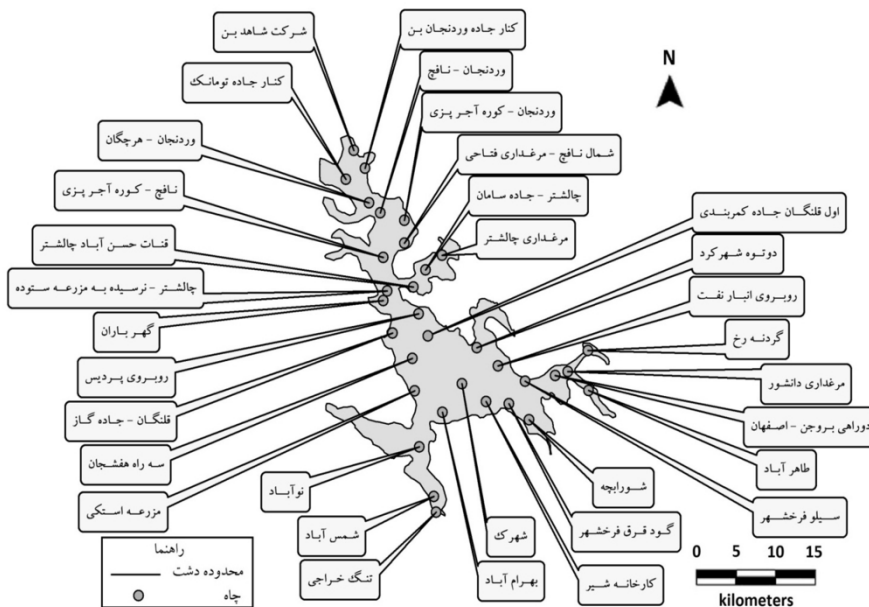
$$Test Error \leq Train Error + Complexity of Set of Models \left( \propto \frac{\|w\|^2}{2} \right) \quad (6)$$

$$Minimize: \frac{\|w\|^2}{2} \quad subject \ to: \begin{cases} w \cdot x_i + b - y_i \leq \epsilon & for \ i = 1, 2, \dots, N \\ y_i - (w \cdot x_i + b) \leq \epsilon & for \ i = 1, 2, \dots, N \end{cases} \quad (7)$$

SVM فقط می‌تواند به صورت خطی به داده‌ها برازش شود و اگر چیدمان داده‌ها در فضای اصلی غیرخطی باشد آنها را به فضای بالاتری می‌برد تا نحوه قرارگیری آنها به صورت خطی شود. از طرفی به دلیل این که یافتن تابع ویژگی برای مسائل مختلف کار بسیار دشواری است بنابراین می‌توان از توابع کرنل<sup>۷</sup> استفاده نمود (Hofmann et al., 2008). بنابراین با حل مسأله بهینه‌سازی،

**ب**

ردیف	نام محل	(متر) $X_{TDM}$	(متر) $Y_{TDM}$	تراز دانه چاه (تراز سطح آب در سال ۱۳۸۸ متر)
۱	اول فلنگان جاده کمربندی	۹۸۳۳۸	۳۵۷۶۸۸	۲۰۳۳۸۲
۲	بهرام آباد	۹۸۳۲۱۹	۳۵۸۰۹۵	۲۰۳۲۷۰
۳	تنگ خراچی	۹۸۳۲۷۷	۳۵۵۳۳۰	۲۰۳۱۰۰
۴	چالشر - جاده سامان	۹۸۲۱۰۵	۳۵۸۹۵۱	۲۰۳۲۳۳
۵	چالشر - زمسیده - مزرعه ستوده	۹۷۹۲۹۹	۳۵۸۱۰۳	۲۰۳۰۷۱۹
۶	دوتوه شهرکرد	۹۷۹۰۹۹	۳۵۵۵۱۲	۲۰۳۵۳۸
۷	دوراهی بروجن - اصفهان	۵۰۱۲۵۵	۳۵۷۹۸۰	۲۰۳۵۳۳
۸	روبروی آزار نفت	۹۸۳۱۳۳	۳۵۷۳۰۸	۲۰۳۳۳۸
۹	روبروی پردیس	۹۸۳۲۳۳	۳۵۷۳۳۳	۲۰۳۰۸۳
۱۰	سه راه هفشجان	۹۸۳۲۶۵	۳۵۷۳۳۳	۲۰۳۵۱۵
۱۱	سیلو فرخشهر	۹۸۳۲۶۶	۳۵۷۳۳۳	۲۰۳۵۱۶
۱۲	شرکت شاهد بن	۹۷۹۱۹۱	۳۵۸۳۲۵	۲۰۳۰۳۶
۱۳	شمال نافع - مرغداری فلاحی	۹۸۱۲۷۷	۳۵۷۵۳۸	۲۰۳۹۰۸
۱۴	شمال نافع - مرغداری فلاحی	۹۸۵۱۳۳	۳۵۵۳۳۳	۲۰۳۰۳۰
۱۵	شهرک	۹۸۳۲۳۳	۳۵۷۳۳۳	۲۰۳۰۳۰
۱۶	نواآباد	۹۷۹۰۰۵	۳۵۷۳۳۳	۲۰۳۰۳۰
۱۷	عطار آباد	۵۰۳۳۰۱	۳۵۷۰۳۳	۲۰۳۰۳۳
۱۸	فلنگان - جاده گاز	۹۸۰۰۱۲	۳۵۷۳۰۳	۲۰۳۰۳۳
۱۹	فلات حسن آباد چالشر	۹۸۳۲۸۸	۳۵۸۳۳۳	۲۰۳۰۳۳
۲۰	کارخانه شیر	۹۸۳۲۶۶	۳۵۷۳۳۳	۲۰۳۰۳۳
۲۱	کنار جاده تومانک	۹۷۹۱۸۸	۳۵۸۳۳۳	۲۰۳۰۳۳
۲۲	کنار جاده وردنجان بن	۹۷۹۲۷۷	۳۵۸۳۳۳	۲۰۳۰۳۳
۲۳	گروه روح	۵۰۳۳۰۱	۳۵۷۳۳۳	۲۰۳۰۳۳
۲۴	گهر باران	۹۷۹۲۶۶	۳۵۸۳۳۳	۲۰۳۰۳۳
۲۵	گود قرق فرخشهر	۹۷۹۲۷۷	۳۵۸۳۳۳	۲۰۳۰۳۳
۲۶	مرغداری چالشر	۹۸۳۲۳۳	۳۵۷۳۳۳	۲۰۳۰۳۳
۲۷	مرغداری دانشور	۵۰۱۲۵۵	۳۵۷۳۳۳	۲۰۳۰۳۳
۲۸	مزرعه استکی	۹۸۳۲۳۳	۳۵۷۳۳۳	۲۰۳۰۳۳
۲۹	نافع - کوره آجر پزی	۹۷۹۲۳۳	۳۵۷۳۳۳	۲۰۳۰۳۳
۳۰	نواآباد	۹۷۹۲۳۳	۳۵۷۳۳۳	۲۰۳۰۳۳
۳۱	وردنجان - کوره آجر پزی	۹۸۱۲۷۷	۳۵۸۳۳۳	۲۰۳۰۳۳
۳۲	وردنجان - نافع	۹۷۹۲۳۳	۳۵۷۳۳۳	۲۰۳۰۳۳
۳۳	وردنجان - هرچگان	۹۷۹۰۰۱	۳۵۷۳۳۳	۲۰۳۰۳۳



شکل ۴- چاه‌های مشاهده‌ای مورد استفاده در مدل‌های LDA، QDA و SVM طی سال‌های ۱۳۷۸ تا ۱۳۸۸ (الف): نقشه پراکنش، (ب): مشخصات

بیشترین افت سطح آب در دوره‌های ۱، ۳، ۵ و ۱۰ ساله به ترتیب برابر با ۷/۹، ۱۱/۹۸، ۴/۳۲ و ۹/۹۴ متر و به همین ترتیب، بیشترین افزایش سطح آب در این دوره‌ها برابر با ۱/۰۵، ۳/۷۵، ۷/۷ و ۱/۰۲ متر بوده است. با وجود این که بیشترین افزایش سطح آب برابر با ۷/۷ متر بوده است ناگزیر به علت رعایت تقارن بازه ۸ متری سایر گروه‌ها (B و C)، گروه A فاصله ۴ تا ۱۲ متر را دربر گرفته است. این موضوع نیز قابل توجه است که چاه‌های درون گروه A با توجه به افزایش سطح آب در این گروه از وضعیت مناسبی برای برداشت برخوردار هستند. با توجه به افت سطح آب در چاه‌های درون گروه‌های B و C لزوم توجه به مدیریت برداشت از این چاه‌ها بارز است. شکل ۵ وضعیت چاه‌های مشاهده‌ای دشت را در دوره‌های زمانی ۱، ۳، ۵ و ۱۰ ساله در گروه‌های A، B و C نشان می‌دهد. ناحیه ای که مدل‌های LDA، QDA و SVM در همه دوره‌های زمانی نتوانسته‌اند گروه‌بندی کنند (ناحیه شرق دشت) جز هر سه گروه بوده است. این نکته قابل توجه است که چاه‌های گروه A در این ناحیه قرار دارند در حالی که چاه‌های گروه‌های B و C علاوه بر این که در این ناحیه قرار دارند در سایر نواحی دشت نیز پراکنده هستند. بنابراین همه مدل‌ها از توانایی کمتری در مدل‌سازی در گروه A برخوردار بوده‌اند.

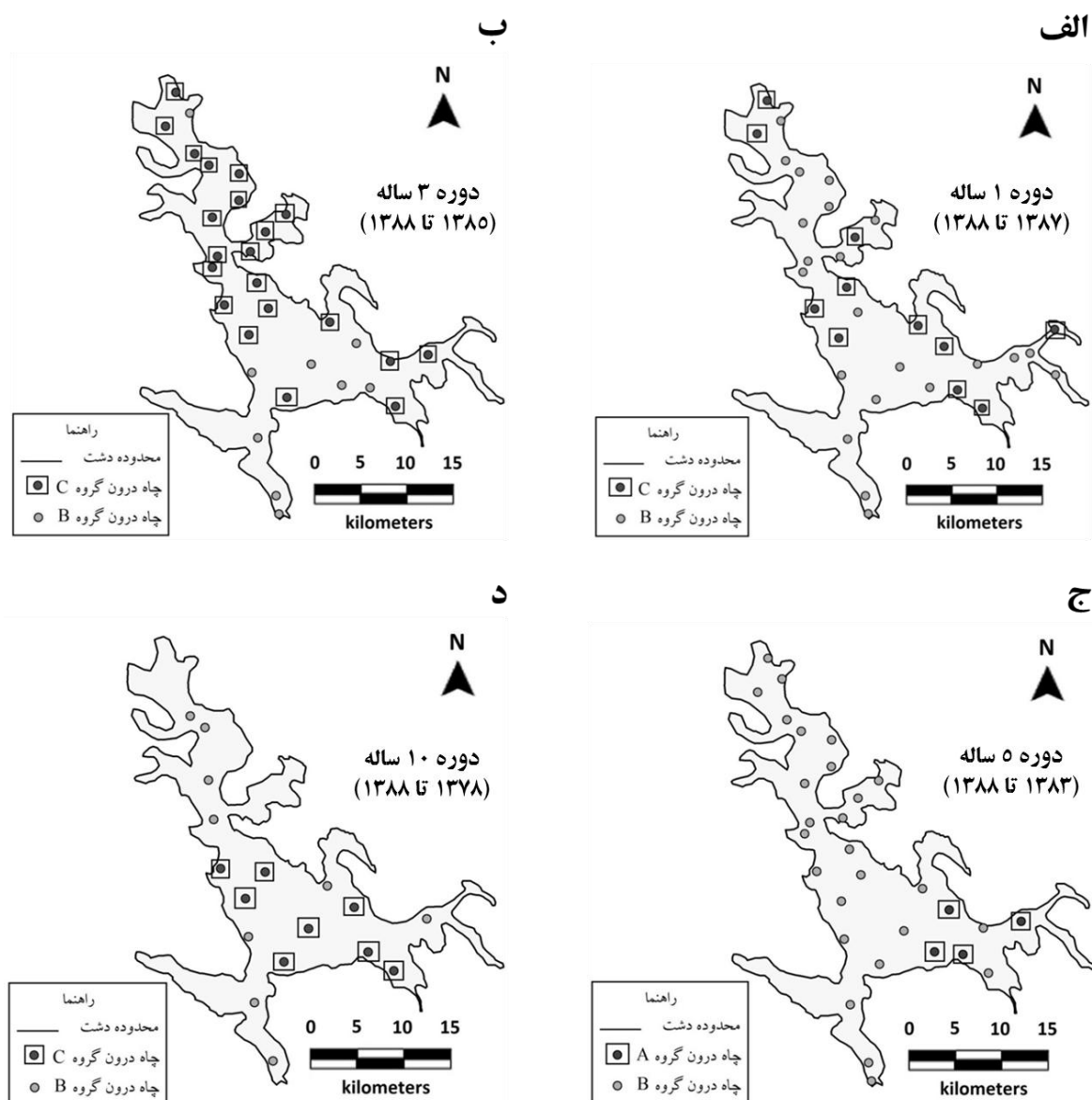
**۲-۳-۱- مدل‌های LDA، QDA و SVM**

از مدل‌های تحلیل تفکیک خطی، تحلیل تفکیک درجه دوم و ماشین

چاه‌های مشاهده‌ای مورد استفاده و ویژگی‌های آنها در شکل ۴ مشخص شده است. به منظور بررسی روند تغییرات سطح آب درون دشت، داده‌های مربوط به آن در سه گروه A، B و C تفکیک شده‌اند. به طوری که داده‌های گروه A مبین افزایش سطح آب در درون دشت (بین ۴ تا ۱۲ متر)، داده‌های گروه B تغییرات بین ۴ متر افت سطح آب تا ۴ متر افزایش آن و داده‌های گروه C مبین افت شدید سطح آب می‌باشد (جدول ۱). انتخاب این گروه‌بندی به این دلیل بوده است که با توجه به داده‌های موجود و رعایت تقارن کلاس‌ها، در سایر گروه‌بندی‌ها کلاس‌های تک عضوی ایجاد می‌شود. این در حالی است که در مدل‌های LDA و QDA حداقل تعداد عضوهای هر گروه دو می‌باشد. شایان ذکر است که تغییرات سطح آب در دشت شهرکرد قابل توجه است که به این مسئله در تحقیقات سایر پژوهشگران نیز اشاره شده است (میرزایی و همکاران، ۱۳۸۵).

جدول ۱- گروه‌بندی بر اساس میزان تغییرات سطح آبخوان در مدل‌های LDA، QDA و SVM در دوره‌های زمانی ۱، ۳، ۵ و ۱۰ ساله طی سال‌های ۱۳۷۸ تا ۱۳۸۸

گروه	میزان تغییرات سطح آب (متر)	تعریف
A	۱۲ تا ۴	افزایش سطح آب
B	۴ تا -۴	افت و افزایش سطح آب
C	-۴ تا -۱۲	افت شدید سطح آب



شکل ۵- نقشه موقعیت چاه‌های مشاهده‌ای در گروه‌های A، B و C برای مدل‌سازی در مدل‌های LDA، SVM و QDA طی سال‌های ۱۳۷۸ تا ۱۳۸۸ (الف): دوره زمانی ۱ ساله، (ب): دوره زمانی ۳ ساله، (ج): دوره زمانی ۵ ساله، (د): دوره زمانی ۱۰ ساله

با دادن مختصات ۲۰ درصد چاه دیگر، توانایی مدل در شبیه‌سازی گروه برای هر چاه مورد آزمون قرار گرفت. به این ترتیب تعداد چاه‌هایی که مدل در هنگام آزمون قادر به قرار دادن در گروه مربوط به آن چاه نبوده است مشخص گردید. همچنین در مدل SVM، تابع کرنل RBF به کار گرفته شده است چرا که تابع RBF به عنوان یکی از بهترین توابع کرنل گزارش شده است (Dibike et al., 2001).

جهت مقایسه و ارزیابی عملکرد مدل‌های مورد بررسی از معیار میانگین خطای نسبی (MRE)<sup>۹</sup> استفاده شده است (رابطه ۱۰).

بردار پشتیبان برای گروه‌بندی چاه‌ها استفاده شده است. به منظور بررسی عملکرد این مدل‌ها، در هر مرتبه تکرار ۸۰ درصد از داده‌ها (تعداد ۲۶، ۲۴، ۲۴ و ۱۴ چاه به ترتیب در دوره‌های ۱ ساله، ۳ ساله، ۵ ساله و ۱۰ ساله) برای آموزش مدل و ۲۰ درصد باقیمانده (تعداد ۷، ۶ و ۳ چاه به ترتیب در دوره‌های ۱ ساله، ۳ ساله، ۵ ساله و ۱۰ ساله) برای آزمون مدل به صورت تصادفی جداسازی شدند. در جهت امکان مقایسه و تحلیل بهتر نتایج مدل، مدل‌سازی برای هر دوره زمانی در ۱۰۰ مرتبه تکرار صورت پذیرفت. در هر تکرار، مختصات و تغییرات سطح آب (گروه A یا B یا C با توجه به جدول ۱) مربوط به ۸۰ درصد از چاه‌ها برای آموزش به مدل معرفی شدند و

$$MRE = \frac{\sum_{i=1}^{100} \left( \frac{x_i}{n_i} \right)}{N} \times 100 \quad (10)$$

که  $x_i$  تعداد چاه گروه‌بندی نشده توسط مدل در مرحله آزمون در تکرار  $i$ ام،  $n_i$  تعداد داده‌های مورد آزمون در تکرار  $i$ ام و  $N$  تعداد کل داده‌های مورد آزمون در ۱۰۰ تکرار می‌باشند.

### ۳- نتایج و بحث

با توجه به جدول‌های ۲ و ۳ مشخص می‌شود که در دوره‌های ۱ و ۳ SVM، در دوره ۵ ساله LDA و در دوره ۱۰ ساله QDA بهترین عملکرد را داشته‌اند. بنابراین ماشین بردار پشتیبان با داشتن کمترین مقدار میانگین خطای نسبی در دو دوره زمانی (۱ و ۳ ساله) و دقت قابل قبول در دو دوره دیگر (۵ و ۱۰ ساله) نسبت به دو روش دیگر برتری داشته است. احتمال می‌رود به دلیل وجود روندی مشخص، در ابتدا با طولانی شدن دوره (از دوره ۱ ساله تا ۵ ساله)، کارایی مدل‌ها افزایش و سپس در دوره ۱۰ ساله کاهش یافته است. همچنین در این تحقیق چاه مشاهده‌ای که مدل، گروه صحیحی برای آن در نظر گرفته است با نام "چاه گروه‌بندی شده" و چاه مشاهده‌ای که مدل به اشتباه در گروه دیگر قرار داده است با نام "چاه گروه‌بندی نشده" معرفی شده است.

#### جدول ۲- میانگین تعداد چاه گروه‌بندی نشده در روش‌های

SVM و QDA، LDA

میانگین تعداد چاه گروه‌بندی نشده در مرحله آزمون (میانگین در ۱۰۰ تکرار)			
دوره زمانی (سال)	مدل LDA	مدل QDA	مدل SVM
۱	۴	۴/۶	۲/۷
۳	۲/۹	۱/۹	۱/۲
۵	۰/۶	۱/۴	۱/۱
۱۰	۱/۲	۰/۴	۱/۲

#### جدول ۳- میانگین خطای نسبی (MRE) در روش‌های LDA،

SVM و QDA

میانگین خطای نسبی (MRE) در مرحله آزمون در ۱۰۰ تکرار (%)			
دوره زمانی (سال)	مدل LDA	مدل QDA	مدل SVM
۱	۸/۱۶	۹/۳۹	۵/۵۱
۳	۸/۰۵	۵/۲۸	۳/۳۴
۵	۱/۶۷	۳/۸۹	۳/۰۵
۱۰	۱۳/۳۴	۴/۴۵	۱۳/۳۴

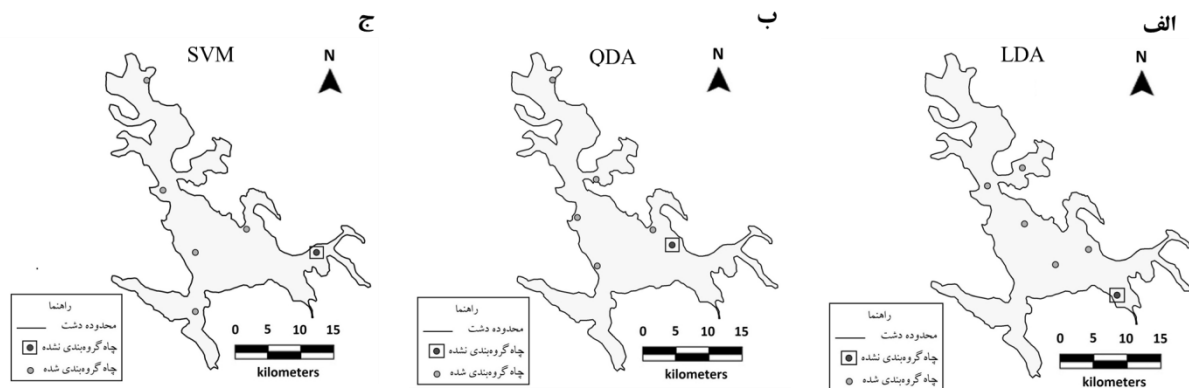
میانگین خطای نسبی مدل‌ها در دوره‌های زمانی ۱، ۳، ۵ و ۱۰ ساله (جدول ۳) نشان دهنده عملکرد مناسب آن‌ها بوده است.

ویژگی کاربردی مدل‌های ماشین بردار پشتیبان، تحلیل تفکیک خطی و درجه دوم این است که می‌توان با در نظر داشتن مختصات هر نقطه از دشت (معرفی مختصات دلخواه در مرحله آزمون) وضعیت تغییرات سطح آب را بدون حفر چاه مشخص نمود و این که نقطه مورد نظر در کدام یک از گروه‌های تغییرات سطح آبخوان قرار خواهد گرفت و در صورتی که محل احداث چاه در گروه B (افزایش یا کاهش سطح آب) قرار گیرد باید لزوم حفر آن مورد بررسی قرار گیرد و در صورت حفر، برداشت از آن مدیریت شود. اگر محلی در گروه A (افزایش سطح آب) باشد حفر چاه در آن مکان ترجیح داده می‌شود. همچنین از حفر چاهی که در گروه C (افت شدید سطح آب) قرار گیرد تا حد امکان خودداری گردد.

به عنوان نمونه، توانایی گروه‌بندی مدل‌های LDA، QDA و SVM در دوره زمانی ۵ ساله در مرحله آزمون در یک تکرار متفاوت (برای LDA تکرار ۳۰ ام، برای QDA تکرار ۶۰ ام و برای SVM تکرار ۹۰ ام) در شکل ۶ مشخص شده است. به منظور امکان قضاوت صحیح در مورد توانایی مدل و همچنین افزایش عامل تصادف، تکرارهای متفاوت به کار رفته است. در این شکل مشاهده می‌شود همه مدل‌ها در شرق دشت توانایی شبیه‌سازی نوسان سطح آبخوان را نداشته‌اند و چاه‌های این ناحیه را صحیح گروه‌بندی نکرده‌اند. این بدان معنی است که مدل‌های فوق از ایجاد ارتباط بین تغییرات سطح آب در شرق دشت و مختصات جغرافیایی آن ناتوان بوده‌اند. به بیان ساده‌تر، تغییرات سطح آب در این ناحیه بر خلاف سایر نواحی دشت از خصوصیات جغرافیایی حوضه بر پایه مختصات زمینی تبعیت نمی‌کند.

#### ۴- نتیجه‌گیری

در مقاله حاضر توانایی روش‌های ماشین بردار پشتیبان (SVM)، تحلیل تفکیک خطی (LDA) و درجه دوم (QDA) در شبیه‌سازی نوسانات دشت شهرکرد مورد تحلیل و بررسی قرار گرفت. میانگین خطای نسبی این مدل‌ها در دوره ۱ ساله به ترتیب برابر با ۵/۵۱، ۸/۱۶ و ۹/۳۹ درصد بوده است. به همین ترتیب میانگین خطای نسبی مدل‌ها در دوره ۳ ساله برابر با ۳/۳۴، ۸/۰۵ و ۵/۲۸ درصد؛ در دوره ۵ ساله برابر با ۳/۰۵، ۱/۶۷ و ۳/۸۹ درصد و در دوره ۱۰ ساله برابر با ۱۳/۳۴، ۴/۴۵ و ۱۳/۳۴ درصد به دست آمده است. نتایج حاکی از این است که ماشین بردار پشتیبان با داشتن کمترین مقدار خطا در دوره‌های ۱ و ۳ ساله و دقت قابل قبول در دوره‌های ۵ و ۱۰ ساله دارای بهترین عملکرد در مدل‌سازی بوده است.



شکل ۶- گروه‌بندی در مرحله آزمون در دوره ۵ ساله (۱۳۸۳ تا ۱۳۸۸): الف: مدل LDA، ب: مدل QDA، ج: مدل SVM

چیت‌سازان م، میرزایی ی، چینی‌پرداز ر (۱۳۸۶) منطقه‌بندی آبخوان دشت شهرکرد با استفاده از تحلیل سری‌های زمانی. مجله علوم دانشگاه شهید چمران، سال جدید، شماره ۱۷: ۱-۱۵.

چینی‌پرداز ر، رکابدار ق، یوسفی‌حاجی‌آباد ر (۱۳۸۵) بررسی توسعه انسانی کشورها با استفاده از روش تحلیل ممیزی آمیخته. فصلنامه بررسی‌های اقتصادی، دوره ۳، شماره ۳: ۵-۲۰.

کوهستانی ن (۱۳۹۰) برآورد سطح آب زیرزمینی با استفاده از سیستم استنتاجی فازی- عصبی، مطالعه موردی: حوزه نرماب استان گلستان. چهارمین کنفرانس مدیریت منابع آب ایران، دانشگاه صنعتی امیرکبیر، تهران، ایران، ۱۳-۱۴ اردیبهشت.

لاله‌زاری ر، طباطبایی س ح (۱۳۸۹) خصوصیات شیمیایی آب زیرزمینی دشت شهرکرد. مجله محیط شناسی، سال ۳۶، شماره ۵۳: ۵۵-۶۲.

محرمپور م، محرابی ع، کاتوزی م (۱۳۹۰) به کارگیری ماشین بردار پشتیبان SVM برای پیش‌بینی دبی روزانه. چهارمین کنفرانس مدیریت منابع آب ایران، دانشگاه صنعتی امیرکبیر، تهران، ایران، ۱۳-۱۴ اردیبهشت.

میرزایی ی، چیت‌سازان م، چینی‌پرداز ر، صمدی ح (۱۳۸۵) پیش‌بینی وضعیت آب‌های زیرزمینی دشت شهرکرد با استفاده از مدل‌های سری‌های زمانی و بررسی راهکارهای بهبود. اولین همایش منطقه‌ای بهره‌برداری بهینه از منابع آب حوضه‌های کارون و زاینده رود، دانشگاه شهرکرد، شهرکرد، ایران، ۱۴-۱۵ شهریور.

Behzad M, Asghari K, Coppola EJr (2010) Comparative study of SVMs and ANNs in aquifer water level prediction. Journal of Computing in Civil Engineering 24(5):408-413.

همه مدل‌ها نیز از توانایی شبیه‌سازی تغییرات سطح آب زیرزمینی در شرق دشت برخوردار نبوده‌اند. علت این امر می‌تواند دخیل نبودن عوامل ژئوفیزیکی در این ناحیه باشد. همچنین قابلیت کاربردی مدل‌های یادشده این است که می‌توان با معرفی مختصات دلخواه در مرحله آزمون، وضعیت تغییرات سطح آب در هر نقطه از دشت را بدون حفر چاه مشخص نمود.

#### پی‌نوشت‌ها

- 1- Support Vector Machine
- 2- Linear Discriminant Analysis
- 3- Quadratic Discriminant Analysis
- 4- Statistical Learning Theory
- 5- Structural Risk Minimization
- 6- Optimum Separating Hyperplane
- 7- Kernel Functions
- 8- Radial Basis Function
- 9- Mean Relative Error

#### ۵- مراجع

باقری‌هارونی م ح، فتحیان ف (۱۳۹۰) بررسی عملکرد شبکه عصبی، نرو- فازی و مدل‌های سری زمانی در پیش‌بینی سطح آب زیرزمینی، منطقه موردی: دشت ابرکوه. اولین همایش منطقه‌ای توسعه منابع آب، دانشگاه آزاد اسلامی واحد ابرکوه، یزد، ایران، ۲۹ اردیبهشت.

پورکرمان م، چرخابی ا، مصدق ح، پیروان ح (۱۳۹۰) شناسایی عوامل شیمیایی تأثیرگذار بر مقاومت خاک با استفاده از آنالیز تفکیکی به روش قدم به قدم، مطالعه موردی: مارن‌های حوضه آبخیز سرخه. مجله علوم و مهندسی آبخیزداری ایران، سال ۵، شماره ۱۶: ۲۳-۳۲.



- meteorological effect implications—A comparative study among soft computing techniques. *Computers & Geosciences* 56:32-44.
- Sudheer Ch, Shrivastava NA, Panigrahi BK, Mathur S (2011) Groundwater level forecasting using SVM-QPSO. In: Proc. Swarm, Evolutionary, and Memetic Computing, Second International Conference, (SEMCCO 2011), 19-21 December, Visakhapatnam, Andhra Pradesh, India, Part I, Lecture Notes in Computer Science 7076:731-741.
- Suryanarayana Ch, Sudheer Ch, Mahmood V, Panigrahi BK (2014) An integrated wavelet-support vector machine for groundwater level prediction in Visakhapatnam, India. *Neurocomputing* 145:324-335.
- Vapnik VN (1998) *Statistical Learning Theory*. John Wiley and Sons, New York, 736p.
- Yoon H, Jun SC, Hyun Y, Bae GO, Lee KK (2011) A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *Journal of Hydrology* 396(1-2):128-138.
- Zhao W, Gao Y, Li C (2012) RVM based on PSO for groundwater level forecasting. *Journal of Computers* 7(5):1073-1079.
- Chen ST, Yu PS (2007) Pruning of support vector networks on flood forecasting. *Journal of Hydrology* 347(1-2):67-78.
- Dibike YB, Velickov S, Solomatine DP, Abbott MB (2001) Model induction with support vector machines: introduction and applications. *Journal of Computing in Civil Engineering* 15(3):208-216.
- Du Z, Jin X (2008) Multiple faults diagnosis for sensors in air handling unit using Fisher discriminant analysis. *Energy Conversion and Management* 49(12):3654-3665.
- Güraksın GE, Haklı H, Uğuz H (2014) Support vector machines classification based on particle swarm optimization for bone age determination. *Applied Soft Computing* 24:597-602.
- Hofmann T, Scholkopf B, Smola AJ (2008) Kernel methods in machine learning. *The Annals of Statistics* 36(3):1171-1220.
- Huberty CJ (1994) *Applied Discriminant Analysis*. John Wiley and Sons, New York, 466p.
- Rani N, Sinha RK, Prasad K, Kedia DK (2011) Assessment of temporal variation in water quality of some important rivers in middle Gangetic plains, India. *Environ Monit Assess* 174(1-4):401-415.
- Shiri J, Kisi O, Yoon H, Lee KK, Nazemi AH (2013) Predicting groundwater level fluctuations with