علوم و مهندسی آبیاری
جلد ۴۱، شماره‌ی ۱، بهار ۹۷، ص. ۱۸۰-۱۶۵

شبیه‌سازی نوسانات سطح آب زیرزمینی با استفاده از ترکیب ماشین بردار پشتیبان و تبدیل موجک

علی اسکندری[۱]، اباذر سُلگی[۲*] و حیدر زارعی[۳]

۱- مربی گروه مهندسی عمران، دانشگاه آزاد اسلامی، واحد بوشهر، ایران.
۲*- نویسنده مسئول، دانشجوی دکتری مهندسی منابع آب، دانشکده مهندسی علوم آب، دانشگاه شهید چمران اهواز abazar_solgi@yahoo.com
۳- دانشیار گروه هیدرولوژی و منابع آب، دانشکده مهندسی علوم آب، دانشگاه شهیدچمران اهواز، ایران.

تاریخ دریافت: ۹۴/۱۰/۱۲ تاریخ پذیرش: ۹۵/۹/۲۰

چکیده

امروزه در بسیاری از کشورهای جهان، به ویژه در مناطقی که با کمبود آب‌های سطحی مواجه هستند، بهره‌برداری از منابع آب زیرزمینی بیش از پیش مورد توجه قرار گرفته است. بهره‌برداری بی‌رویه از این منابع، بدون بهره‌گیری از مطالعات منابع آب زیرزمینی می‌تواند مشکلات و پیامدهای جبران‌ناپذیری را به‌بار آورد. مدیریت صحیح این منابع با شناخت کامل و آگاهی از این منابع امکان‌پذیر است. در این تحقیق از مدل ماشین بردار پشتیبان (SVM) برای پیش‌بینی ماهانه سطح آب زیرزمینی دشت برازجان در استان بوشهر در یک دوره ۱۲ ساله (۱۳۹۲-۱۳۸۱) استفاده شد. به منظور بهبود نتایج این مدل، ابتدا داده‌های ورودی با استفاده از تبدیل موجک به زیر سیگنال‌هایی تبدیل شدند. سپس زیرسیگنال‌های مزبور، به‌عنوان ورودی به مدل SVM به‌عنوان مدل ترکیبی ماشین بردار پشتیبان-موجک (WSVM) استفاده شدند. برای پیش‌بینی سطح آب زیرزمینی از پنج حلقه چاه مشاهده‌ای به‌همراه متغیرهای سطح آب زیرزمینی، بارش، تبخیر و دما، استفاده شد. نتایج نشان داد مدل ترکیبی WSVM دارای عملکرد بهتری نسبت به مدل SVM می‌باشد. به‌طوری که عملکرد مدل ترکیبی در شبیه‌سازی ماهانه سطح آب زیرزمینی ۸ درصد نسبت به مدل SVM بهتر بود. این روش ایده‌ای جدید برای افزایش دقت مدل‌ها و کاربرد آن‌ها در مسائل هیدروژئولوژی می‌باشد. در پایان سطح آب زیرزمینی برای یک سال آینده با مدل ترکیبی برآورد گردید.

کلید واژه‌ها: ماشین بردار پشتیبان، تبدیل موجک، سطح آب زیرزمینی، شبیه‌سازی.

مقدمه

منابع آب‌های زیرزمینی یکی از مهم‌ترین و با ارزش‌ترین منابع آب به‌شمار می‌روند. شناخت صحیح و بهره‌برداری اصولی از آن‌ها به‌خصوص در مناطق خشک و نیمه خشک در توسعه پایدار بسیاری از فعالیت‌های کشاورزی، اجتماعی و اقتصادی تأثیر به‌سزایی دارد. تاکنون مدل‌های زیادی برای پیش‌بینی سطح آب زیرزمینی به‌کار برده شده است. از جمله این مدل‌ها می‌توان به مدل‌های سری زمانی تجربی و مدل‌های فیزیکی اشاره کرد (Izadi et al., 2007). علیرغم استفاده گسترده از این مدل‌ها، در شرایط رفتار دینامیکی سامانه هیدرولوژیکی، مدل‌های مزبور توانایی کافی را برای تخمین پارامترهای مورد نیاز در شبیه‌سازی سطح آب زیرزمینی ندارند (Bierkens, 1998). علاوه بر این، مدل‌های فیزیکی در عمل به داده‌های زیادی برای شبیه‌سازی نوسانات سطح آب زیرزمینی نیاز دارند. از طرفی به‌دلیل غیرخطی بودن روابط بین متغیرهای مؤثر بر سطح آب زیرزمینی، مدل‌های فیزیکی در ارائه رابطه بین متغیرها، فاقد توانایی لازم هستند (Nayak, Satyaji Rao and Sudheer, 2006).

Wang و Ding (2003) از ترکیب تحلیل موجک و شبکه عصبی مصنوعی برای پیش‌بینی تراز آب زیرزمینی شهر پکن و دبی رودخانه یانگ تسه استفاده کردند. Nourani، Asghari Mogaddam و Naderi (2008) مدلی بر پایه شبکه عصبی مصنوعی برای پیش‌بینی فضایی زمانی (Spatiotemporal) سطح آب زیرزمینی در آبخوان چند لایه تبریز ارائه دادند. راندمان مدل STANN (Temporal-ANN Spatio) با دو مدل ترکیبی NG (Neural-Geostatistics) و TSG (Multivariate Time Series-Geostatistics) مقایسه شد. نتایج نشان داد که مدل STANN نتایج قابل قبولی ارائه داد. Adamowski و Chan (2011) از روش ترکیبی شبکه عصبی مصنوعی-موجک (Wavelet-Artificial Neural Network) برای پیش‌بینی نوسانات ماهانه سطح آب زیرزمینی استفاده کردند. آن‌ها نتایج حاصل از مدل ترکیبی مزبور را با مدل های شبکه عصبی مصنوعی (Artificial Neural Network) و میانگین متحرک یک‌پارچه اتورگرسیو (Auto Regressive Integrated Moving Average) مقایسه نمودند. نتایج نشان داد که مدل WANN، برای پیش‌بینی سطح آب زیرزمینی دقیق‌تر از هر یک از مدل‌های ANN و ARIMA می‌باشد. Nakhaei، Saberi nasr و Faraj zadeh (2011) مدل ترکیبی شبکه عصبی مصنوعی-موجکی را برای پیش‌بینی نوسانات سطح آب

زیرزمینی دشت قروه مورد استفاده قرار دادند. نتایج نشان داد که مدل شبکه عصبی مصنوعی-موجک نتایج بهتری از مدل عددی MODFLOW داشته و می‌تواند در زمینه پیش‌بینی سطح آب جایگزین MODFLOW شود.

Fallah-Mehdipour، Bozorg Haddad و Mariño (2013) برای شبیه‌سازی سطح آب زیرزمینی از برنامه‌ریزی ژنتیک (Genetic Programming) استفاده نمودند. آن‌ها نتایج مدل برنامه‌ریزی ژنتیک را با مدل سیستم استنتاج فازی عصبی تطبیقی(Adaptive Neuro-Fuzzy Inference System) مقایسه نمودند. نتایج نشان داد که مدل GP، یک ابزار مؤثر در شبیه‌سازی سطح آب زیرزمینی است. Shiri et al. (2013) برای پیش‌بینی نوسانات سطح آب زیرزمینی با بررسی اثرات هواشناسی، از روش‌های مختلف ANN، ANFIS، ماشین بردار پشتیبان (Support Vector Machine) و برنامه‌ریزی بیان ژن (Gene Expression Programming) استفاده نمودند. در این بررسی از داده‌های سطح آب زیرزمینی، بارش و مقدار تبخیر در یک دوره هشت ساله استفاده شد. نتایج نشان داد که مدل GEP عملکرد بهتری نسبت به دو روش دیگر داشت. Suryanarayana و Paresh Chandra (2014) از ترکیب رگرسیون‌بردار پشتیبان یکپارچه-موجک (Wavelet-Support Vector Regression) برای پیش‌بینی نوسانات ماهانه سطح آب زیرزمینی استفاده کردند. آن‌ها از داده‌های ماهانه بارش، متوسط دما، حداکثر دما و عمق آب زیرزمینی برای یک دوره ۱۳ ساله استفاده نمودند. مدل‌های WA-SVR، ANN، SVR و ARIMA با هم مقایسه شدند. نتایج نشان داد که مدل WA-SVR نسبت به سایر مدل‌ها در پیش‌بینی سطح آب زیرزمینی عملکرد بهتری دارد.

Moosavi et al. (2014) بهینه‌سازی مدل‌های ترکیبی سیستم استنتاج فازی عصبی تطبیقی-موجک (Wavelet-ANFIS) و شبکه عصبی مصنوعی-موجک (WANN) به‌وسیله روش تاگوچی را برای پیش‌بینی سطح آب زیرزمینی در مشهد انجام دادند. در این تحقیق ساختارهای مختلف برای هر دو مدل ترکیبی مورد بررسی قرار گرفت. نتایج نشان داد که مدل WANFIS عملکرد بهتری نسبت به مدل WANN دارد.

Rajaee و Zeynivand (2015) مدل‌سازی تراز آب زیرزمینی را با بهره‌گیری از مدل ترکیبی شبکه عصبی مصنوعی-موجک برای دشت شریف‌آباد استان قم مورد بررسی قرار دادند. نتایج نشان داد این مدل تطابق خوبی با مقادیر مشاهداتی داشته است.

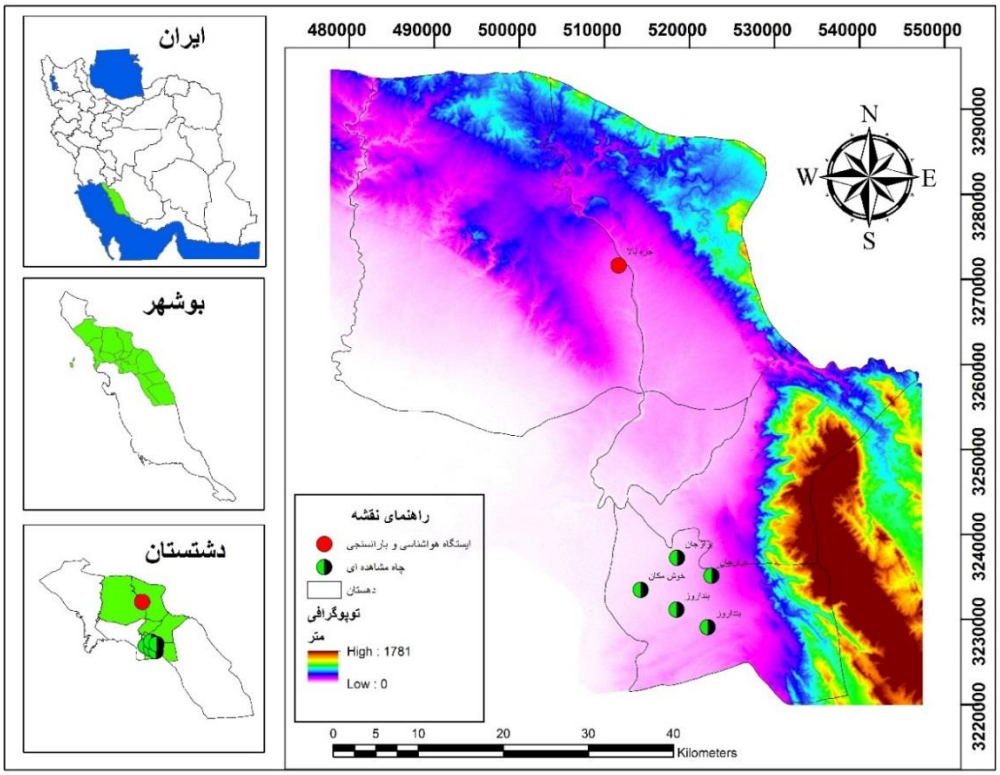
در سال‌های اخیر، استفاده از روش‌های غیرخطی توسعه یافته و با ترکیب تکنیک‌های مختلف برای تشکیل یک ابزار هیبریدی، به بهبود دقت این روش‌ها برای کاربردهای مختلف پرداخته شده است. در سال‌های اخیر تحقیقات خوبی در زمینه استفاده از ANN و ANFIS صورت گرفته است این درحالی است که در بررسی مرور منابع انجام شده، در زمینه‌ی استفاده ترکیبی از SVM گزارشی مشاهده نشده است. لذا در تحقیق حاضر علاوه بر استفاده از مدل SVM، سعی گردید با کمک تبدیل موجک، مدل مناسبی برای پیش‌بینی سطح آب زیرزمینی استفاده نمود.

مواد و روش‌ها

منطقه مورد مطالعه

شهرستان دشتستان در شرق استان بوشهر از جنوب به شهرستان دشتی، از مغرب به شهرستان بوشهر و تنگستان از شمال غرب به شهرستان گناوه و از مشرق و شمال شرق به استان فارس منتهی می‌شود. این شهرستان یکی از قطب‌های کشاورزی استان بوشهر است که حدود چهار میلیون و پانصد نفر نخل در این شهرستان وجود دارد. خرما به‌عنوان مهم‌ترین محصول کشاورزی در این شهرستان محسوب می‌شود. شیوه آبیاری نخلستان‌های دشتستان، عمدتاً به‌صورت غرقابی است. این روش آبیاری باعث کاهش راندمان آب شده، در نتیجه آب بسیار زیادی برای آبیاری نخیلات این شهرستان مصرف می‌شود. یکی از مشکلات مهم در این شهرستان، حفر چاه‌های متعدد و غیرمجاز است، به طوری‌که اضافه برداشت چاه‌ها باعث افت سطح آب زیرزمینی گردیده و دشت‌های این شهرستان را با مشکلات متعددی مواجه ساخته است.

شکل (۱) شهرستان دشتستان و موقعیت چاه‌های مشاهده‌ای در این تحقیق را نشان می‌دهد. جدول (۱) پارامترهای ماهانه بارش، دما، تبخیر و سطح آب زیرزمینی در یک دوره زمانی ۱۲ ساله (۱۳۸۱-۱۳۹۲) ارائه شده است. همچنین در جدول (۲) مشخصات چاه‌های مورد استفاده آورده شده است.



شکل ۱- موقعیت چاه‌های دشتستان در استان بوشهر و ایران

جدول ۱- متغیرهای ماهانه هواشناسی ایستگاه جره بالا در بازه زمانی (۱۳۸۱-۱۳۹۲)

نام ایستگاه	کد ایستگاه	متغیر اقلیمی	میانگین	حداکثر	حداقل	مختصات جغرافیایی (متر)		
						طول	عرض	ارتفاع
جره بالا	۲۳-۰۱۱	بارش (میلی‌متر)	۲۱/۰۶	۲۵۰/۰۰	۰	۵۱۱۶۷۶	۳۲۷۱۶۳۷	۱۰۶
		دما (درجه سانتی‌گراد)	۲۵/۴۶	۳۷/۰۰	۰			
		تبخیر (میلی‌متر)	۳۱۷/۷۲	۷۵۴/۲۰	۰			

جدول ۲- مشخصات چاه‌های مورد استفاده و داده‌های سطح آب زیرزمینی آن‌ها

چاه	نام چاه	تراز سطح ایستابی (متر)			طول جغرافیایی (متر)	عرض جغرافیایی (متر)	ارتفاع از سطح دریا (متر)
		میانگین	حداکثر	حداقل			
۱	خوش مکان	۴/۹۳	۶	۰	۵۱۴۲۸۰	۳۲۳۳۴۹۳	۳۴/۴۵
۲	بنداروز	۹/۸۰	۱۱/۹۸	۰	۵۱۸۴۹۰	۳۲۳۱۱۶۰	۵۰/۰۳
۳	بنداروز	۲۶/۳۳	۲۹/۷۱	۰	۵۲۲۱۶۰	۳۲۲۹۱۰۰	۷۲/۱۹
۴	برازجان	۶/۵۶	۹/۵۶	۰	۵۱۸۵۴۷	۳۲۳۷۲۵۴	۴۴/۵۹
۵	برازجان	۵۶/۳۰	۵۹/۸۹	۰	۵۲۲۶۰۵	۳۲۳۵۱۵۸	۹۸/۵۹

آماده‌سازی داده‌ها

به‌علت اینکه وارد کردن داده‌ها به صورت خام باعث کاهش سرعت و دقت می‌شود از روش استانداردسازی داده‌ها استفاده شده است با روش استانداردسازی هر عدد تبدیل به عددی بین صفر تا یک می‌شود (Riad et al., 2004). با توجه به پیشنهاد Solgi (2014) از رابطه زیر برای نرمال‌سازی (استانداردسازی) استفاده شده است.

$$y = 0.5 + (0.5 \times \left(\frac{X-\bar{X}}{X_{max}-X_{min}}\right)) \qquad (۱)$$

که در این روابط X داده مورد نظر، $\bar{X}$ میانگین داده‌ها، X_{max}

حداکثر داده‌ها، X_{min} حداقل داده‌ها و y داده استاندارد شده می‌باشد. در این تحقیق از ۷۵ درصد داده‌های سطح آب زیرزمینی برای آموزش و ۲۵ درصد برای شبیه‌سازی استفاده شده است.

ماشین بردار پشتیبان (SVM)

ماشین بردار پشتیبان، یکی از روش‌های یادگیری با نظارت (Supervised learning) است. ماشین بردار پشتیبان، دسته‌بندی کننده‌ای است که جزء شاخه روش‌های کرنلی (Kernel Methods) در یادگیری ماشین محسوب می‌شود. این نوع سیستم یادگیری هم برای دسته‌بندی و هم برای تخمین و برآورد تابع برازش داده‌ها به‌کار می‌رود. به‌طوری‌که کمترین خطا در دسته‌بندی داده‌ها و یا تابع برازش رخ دهد. این روش از جمله روش‌های نسبتاً جدیدی است که در سال‌های اخیر کارایی خوبی نسبت به روش‌های قدیمی‌تر برای طبقه‌بندی از جمله شبکه‌های عصبی پرسپترون، داشته است. مبنای کاری دسته‌بندی کننده ماشین بردار پشتیبان، دسته‌بندی خطی داده‌ها است. در دسته‌بندی خطی داده‌ها سعی می‌شود خطی انتخاب گردد که حاشیه اطمینان بیشتری داشته باشد. حل معادلهٔ پیدا کردن خط بهینه برای داده‌ها به‌وسیله روش‌های QP که روش‌های شناخته شده‌ای در حل مسائل محدودیت‌دار هستند صورت می‌گیرد. قبل از تقسیم خطی، برای اینکه ماشین بتواند داده‌های با پیچیدگی بالا را دسته‌بندی کند، داده‌ها به‌وسیله تابع phi به فضای با ابعاد خیلی بالاتر (High dimensional space) برده می‌شود. به‌منظور حل مسأله با ابعاد بالا از قضیه دوگانی لاگرانژ (Lagrange Duality Theorems) استفاده می‌شود. از این رو تابع هدف مربوطه به‌جای تابع پیچیده phi با ابعاد بالا، از تابع ساده‌تری به نام تابع هسته که ضرب برداری تابع phi می‌باشد، استفاده می‌شود.

ماشین بردار پشتیبان در سال ۱۹۹۲ توسط وپنیک (Vapnik) ریاضیدان روس، بر پایه نظریه یادگیری آماری (Statistical Learning Theory) بنا گردید. هدف ماشین بردار پشتیبان، تشخیص تابع f(x) برای الگوهای آموزشی x است، بهطوری‌که حداکثر حاشیه را از مقادیر آموزشی y داشته باشد. به عبارت دیگر، ماشین بردار پشتیبان مدلی است که یک منحنی با ضخامت مشخص به داده‌ها برازش می‌دهد، به‌طوری‌که کمترین خطا در داده‌های آزمایشی صورت گیرد. (Cortes and Vapnik, 1995).

ماشین بردار پشتیبان غیرخطی

Cortes و Vapnik (1995) با استفاده از مفهوم حاصل-ضرب داخلی در فضاهای هیلبرت (Hilbert) و قضیه هیلبرت–اشمیت (Hilbert-Schmidt Theory) نشان دادند که می‌توان بردار ورودی x را با یک تبدیل غیرخطی، به فضایی با بعد بالا، انتقال داد. سپس در فضای مزبور، حاصل‌ضرب داخلی را انجام دادند. آنها ثابت کردند که اگر یک هسته متقارن، شرایط قضیه مرکر(Mercer) را داشته باشد، اعمال هسته مزبور در فضای ورودی با بعد کم، می‌تواند به‌عنوان حاصل‌ضرب داخلی در فضای هیلبرت با بعد زیاد تلقی شود. بدین ترتیب محاسبات به‌طور قابل ملاحظه‌ای کاهش می یابد (Cortes and Vapnik, 1995).

برخی فرم های تابع کرنل، به شرح مندرج در رابطه (۲) می‌باشند:

$$\begin{gathered} k(x,y) = (x.y+1)^p \quad p=2,3,\dots \\ k(x,y) = \exp(-\frac{\|x-y\|^2}{2\sigma^2}) \\ k(x,y) = \tanh(x.y+\theta) \\ k(x,y) = (x.y) \end{gathered} \qquad (۲)$$

به‌ترتیب شامل کرنل چندجمله‌ای (Polynomial)، کرنل گوسی یا RBF، کرنل سیگموئید (Sigmoidal) و کرنل خطی (Linear) می‌باشند (Sujay Raghavendra, and Paresh Chandra, 2014).

کرنل گوسی صرفاً دارای پارامتر g می‌باشد. برای این پارامتر، ابتدا مقدار پیش فرض ۱/k (k تعداد ویژگی‌ها) استفاده می‌شود. سپس مقادیر $1/k^2$ و ۱ در نظر گرفته می‌شوند. در مورد کرنل سیگموئید با توجه به هم‌گرایی به جواب، صرفاً پیش فرض مقدار ثابت صفر و گاما ۱/k استفاده می‌شود. شایان ذکر است که کرنل سیگموئید فاقد پارامتر است. کرنل خطی نیز فاقد پارامتر می‌باشد. لذا تنها یک حالت اجرا دارد. کرنل چندجمله‌ای دارای دو پارامتر است. پارامتر اول درجه چندجمله‌ای است و با d مشخص می‌شود، به ازای مقادیر ۱، ۲، ۳ و ۴ مسائل حل می‌شوند. برای پارامتر دوم نیز که عدد ثابت جمع شونده است و با r مشخص می‌شود، سه حالت ۰، ۱– و۱+ استفاده می‌شود. مسأله بهینه‌سازی دوگان در حالت جدایی‌ناپذیر و غیرخطی به صورت زیر خواهد بود:

$$\begin{cases} Minimize_{\ \alpha} & -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j k(x_i,x_j) \\ & +\sum_{i=1}^{N}\alpha_i \\ Subject\ to & 0 \le \alpha_i \le C \\ & \sum_{i=1}^{N}\alpha_i y_j = 0 \end{cases} \qquad (۳)$$

بردارهای پشتیبان الگوهایی هستند که ضرایب لاگرانژ متناظر آن‌ها در رابطه $0 \le \alpha_i \le C$ صدق کند و تعداد آن‌ها N_b است. برای محاسبه b از رابطه (۴) استفاده می‌شود:

$$\begin{gathered} b_j = y_j - \sum_{i=1}^{N_{SV}}\alpha_i y_i k(SV_i, SV_j) \\ b = \frac{1}{N_b}\sum_{j=1}^{N_b} b_j \end{gathered} \qquad (۴)$$

تابع تصمیم‌گیری مطابق رابطه (۵) می‌باشد:

$$f(x) = Sign(\sum_{i=1}^{N_{SV}} \alpha_i y_i\, k(x, SV_j) + b) \qquad (۵)$$

برای کسب اطلاعات بیشتر به منابع (Schölkopf and Smola, 2002; Smola and Schölkopf, 2004; Schölkopf, Smola and Müller, 2005) مراجعه شود.

تبدیل موجک

تابع موجک، تابعی است که دو ویژگی مهم یعنی نوسانی بودن و کوتاه مدت بودن را دارا باشد. $\Psi(x)$، تابع موجک است اگر و فقط اگر تبدیل فوریه آن $\Psi(\omega)$، شرط ارائه شده در رابطه (۶) را ارضا کند.

$$C_g = \int_{-\infty}^{+\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty \qquad (۶)$$

این شرط به‌عنوان شرط پذیرفتگی برای موجک $\Psi(x)$ شناخته می‌شود. رابطه فوق را می‌توان معادل با فرمول زیر دانست. به‌عبارت دیگر برای اینکه موجک شرط بالا را داشته باشد، باید معادله (۷) برقرار باشد:

$$\int_{-\infty}^{+\infty} \Psi(x) dx = 0 \qquad (۷)$$

این ویژگی تابع با میانگین صفر، چندان محدودکننده نبوده و توابع بسیاری را می‌توان براساس آن، تابع موجک نامید. $\Psi(x)$ تابع موجک مادر است که توابع مورد استفاده در تحلیل با دو عمل ریاضی انتقال (Translation) و مقیاس (Scale) در طول سیگنال مورد مطالعه، تغییر اندازه و محل می‌یابند. در نهایت ضرایب موجک در هر نقطه از سیگنال (مقدار b) و برای هر مقدار از مقیاس (مقدار a) با رابطه (۸) و (۹) قابل محاسبه است.

$$\Psi_{a,b}(x) = \frac{1}{\sqrt{a}} \Psi\left(\frac{x-b}{a}\right) \qquad (۸)$$

$$f(x)\Psi\left(\frac{x-b}{a}\right) dx = \int_{-\infty}^{+\infty} f(x)\Psi_{a,b}(x) dx \qquad (۹)$$

برای تبدیل مقیاس به فرکانس می‌توان از رابطه (۱۰) استفاده نمود.

$$F_a = \frac{F_C}{a.\Delta} \qquad (۱۰)$$

که در آن F_a فرکانس متناظر با مقیاس a و F_C فرکانس اصلی تابع موجک (که فرکانس مرکزی آن موجک نامیده می‌شود) و Δ دوره نمونه‌برداری است. اما در تبدیل موجک هنگام استفاده از یک تابع، خاصیت محو شوندگی ممان‌های آن، از اهمیت زیادی برخوردار است. k امین ممان یک موجک $\Psi(x)$ برابر است با:

$$\int_{-\infty}^{+\infty} x^K \Psi(x) dx = 0 \qquad (۱۱)$$

بنابراین یک موجک در صورتی که رابطه زیر در آن صدق کند، دارای m ممان محوشدنی (Vanishing Moments) می باشد.

$$\int_{-\infty}^{+\infty} x^K \Psi(x) dx = 0 \quad k= 0, 1, 2, \ldots, m\text{-}1 \qquad (۱۲)$$

رابطه (۱۲) نشان می‌دهد که هر تابع موجک حداقل دارای یک ممان محو شونده است.

در تحلیل سیگنال از فرم گسسته موجک (Discrete Wavelet Transform) نیز استفاده می‌شود. درDWT پارامترهای انتقال و مقیاس به طور گسسته انتخاب می‌شوند. به‌طوری‌که:

$$a = 2^{-j}\,,\ b = 2^{-j}k \qquad (۱۳)$$

که j و k اعداد صحیح هستند. در نتیجه با جایگذاری به جای a و b رابطه زیر حاصل می‌شود

$$\Psi_{j,k}(x) = 2^{\frac{j}{2}} \Psi(2^j x - k) \qquad (۱۴)$$

با توجه به رابطه مشخص کننده یک تابع موجک، می‌توان دریافت که توابع بسیاری وجود دارند که دارای این ویژگی‌ها باشند. در سال‌های اخیر، تعداد زیادی توابع موجک گسترش داده شده‌اند که در موارد متعدد، قابلیت‌های گوناگونی دارند. پس از انتخاب تابع موجک مادر، محاسبه ضرایب موجک به ازای هر مقداری (اعداد صحیح و اعشاری مثبت از a مقیاس) قابل انجام است. برای مقاصد گوناگون، محاسبه ضرایب موجک به ازای مقادیر خاصی از a، کفایت می‌کند.

وجود قابلیت‌های نرم‌افزاری برای انجام، سبب شده است که به‌راحتی امکان محاسبه ضرایب موجک برای دامنه وسیعی از a فراهم آید. بدین ترتیب دامنه‌ای از a که برآورنده یک منظور خاص باشد، انتخاب می‌شود-(Kumar, and Fofola Georgiou, 1994; Mallat,1998)

ساختارهای مختلف مدل‌سازی

برای اجرای مدل ماشین بردار پشتیبان با توجه به تعداد پارامترهای ورودی، ۷ نوع ترکیب مختلف مطابق جدول (۳) مورد بررسی قرار گرفت. پارامترهای H_t, P_t, T_t, E_t به ترتیب تبخیر، دما، بارش و تراز سطح ایستابی در دوره زمانی ماهانه و $H_{t-2}, H_{t-1}, P_{t-2}, P_{t-1}, T_{t-2}, T_{t-1}, E_{t-2}, E_{t-1}$ به ترتیب تبخیر، دما، بارش و تراز سطح ایستابی در دوره‌های زمانی گذشته و H_{t+1} تراز سطح ایستابی در دوره آتی می‌باشد.

برای دستیابی به نتایج حاصل از مدل ترکیبی، ابتدا با استفاده از تبدیل موجک، پارامترهای ورودی به زیرسیگنال‌ها تجزیه می گردد. سپس زیرسیگنال‌های مزبور، به عنوان ورودی به مدل ماشین بردار پشتیبان وارد می‌شوند. و بدین ترتیب مدل ترکیبی ماشین بردار پشتیبان–موجک حاصل می‌گردد. یکی از نکات مهم و اساسی در مورد انتخاب موجک‌های مادر، ساختار پدیده و نوع سری زمانی آن می‌باشد. لذا الگویی‌هایی از توابع موجک مادر که به لحاظ شکل هندسی، به سری زمانی منطبق باشند، می‌توانند عمل نگاشت را انجام دهند. بنابراین در تحقیق حاضر، سه تابع موجک زیر انتخاب شدند. شکل (۲) نشان دهنده سه تابع موجک Coif1، Db2 و Db4 می‌باشد. شایان ذکر است برنامه برای توابع موجک مذکور، با درجات مختلف تجزیه و برای کرنل‌های مختلف اجرا شد. به منظور بدست آوردن درجه تجزیه مناسب برای توابع موجک از رابطه (۱۵) به عنوان پیشنهاد اولیه استفاده گردید (Nourani, Komasi and Mano., 2009).

$$L=Int[\log(N)] \qquad (۱۵)$$

در این رابطه L درجه تجزیه پیشنهادی و N تعداد سری زمانی اعداد می‌باشد. در این پژوهش با N=۱۳۲ و L=۲ می‌باشد که برای دقت بیشتر، درجات تجزیه ۱ تا ۴ مورد بررسی قرار گرفت.

از پارامترهای ضریب تعیین، جذر میانگین مربعات خطا و معیار آکائیک در ارزیابی کارایی روش‌ها استفاده شده است که روابط آن‌ها به صورت زیر می‌باشد.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(H_i - H_o)^2}{N}} \qquad (۱۶)$$

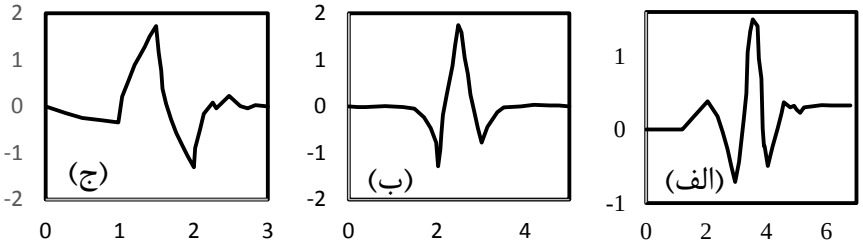
$$R^2 = 1 - \frac{\sum_{i=1}^{N}(H_o - H)^2}{\sum_{i=1}^{N}(H_o - \bar{H})^2} \qquad (۱۷)$$

$$AIC = m \times \ln(RMSE) + 2(Npar) \qquad (۱۸)$$

در این روابط پارامترها عبارتند از N: تعداد داده‌ها، H_0 :داده های مشاهداتی، $\bar{H}$: میانگین داده‌های مشاهداتی، H_i: داده‌های محاسباتی، m: تعداد پارامترها، Npar: تعداد داده‌های آموزش دیده می‌باشد.

جدول ۳- جزئیات ترکیب‌های مختلف مدل ماشین بردار پشتیبان

ترکیب	ورودی	خروجی
۱	H_{t-2}, H_{t-1}, H_t	H_{t+1}
۲	P_{t-2}, P_{t-1}, P_t	H_{t+1}
۳	T_{t-2}, T_{t-1}, T_t	H_{t+1}
۴	E_{t-2}, E_{t-1}, E_t	H_{t+1}
۵	H_t, P_t, T_t, E_t	H_{t+1}
۶	$H_{t-1}, H_t, P_{t-1}, P_t, T_{t-1}, T_t, E_{t-1}, E_t$	H_{t+1}
۷	$H_{t-2}, H_{t-1}, H_t, P_{t-2}, P_{t-1}, P_t, T_{t-2}, T_{t-1}, T_t, E_{t-2}, E_{t-1}, E_t$	H_{t+1}



شکل ۲ – الف) تابع موجک db4 ب) تابع موجک coif1 ج) تابع موجک db2

جدول ٤- نتایج بهترین ساختار در ترکیب‌های مختلف مدل SVM برای چاه یک

ترکیب	پارامترهای مدل					ضریب همبستگی		جذر میانگین مربعات خطا	
	کرنل	Gama	Sin2	درجه	B	آموزش	شبیه‌سازی	آموزش	شبیه‌سازی
۱	Lin	۵/۶۴۷	----	----	۰/۰۰۰۱	۰/۸۸۵	۰/۸۰۶	۰/۰۴۲۱	۰/۰۵۴۱
۲	Poly	۰/۰۰۲	۲۸/۸۲	۳	-۰/۰۱۱۴	۰/۴۶۴	۰/۵۱۸	۰/۰۸۰۲	۰/۲۲۸۱
۳	RBF	۰/۱۸۵	۴/۶۷۶	----	۰/۰۹۵۲	۰/۴۴۶	۰/۵۴۴	۰/۰۸۱۳	۰/۲۲۷۵
۴	Lin	۰/۳۰۳	----	----	۰/۰۰۰۰۱	۰/۴۶۷	۰/۵۰۳	۰/۰۸۰۰	۰/۲۲۵۰
۵	Lin	۰/۲۸۱	----	----	-۰/۰۰۰۱	۰/۹۰۶	۰/۸۵۰	۰/۰۳۸۳	۰/۰۵۳۹
۶	RBF	۱۸/۲۸	۷۸/۷۸	----	۰/۳۲۸۴	۰/۹۱۹	۰/۸۷۲	۰/۰۳۵۷	۰/۰۵۵۰
۷	Poly	۰/۰۰۱	۴۸۴/۳	۳	-۱/۰۴۶۸	۰/۹۱۹	۰/۸۵۹	۰/۰۳۵۷	۰/۰۴۱۴

جدول ۵- نتایج بهترین ساختار در ترکیب‌های مختلف مدل SVM برای چاه‌های ۲-۵

چاه	ترکیب	پارامترهای مدل					ضریب همبستگی		جذر میانگین مربعات خطا	
		کرنل	Gama	Sin2	درجه	B	آموزش	شبیه‌سازی	آموزش	شبیه‌سازی
۲	۶	RBF	۸۱۴/۹۷	۸۴۱/۹	----	۳/۷۲۰	۰/۹۷۲	۰/۹۵۶	۰/۰۱۸۷	۰/۰۱۷۳
۳	۵	Lin	۱/۵۷۸	----	----	-۰/۰۰۰۱	۰/۹۶۰	۰/۸۷۷	۰/۰۲۵۰	۰/۰۲۶۲
۴	۶	Lin	۰/۳۱۷۴	----	----	۰/۰۰۰۰۳	۰/۹۰۴	۰/۸۵۷	۰/۰۳۲۱	۰/۰۴۶۰
۵	۶	RBF	۳۴۹/۷۷	۱۴۶۸	----	۱/۴۸۴۱	۰/۹۶۳	۰/۹۱۴	۰/۰۳۰۷	۰/۰۱۳۱

نتایج و بحث

نتایج مدل‌سازی با مدل ماشین بردار پشتیبان

برای مدل‌سازی با مدل ماشین بردار پشتیبان از نرم‌افزار متلب استفاده گردید. ساختارهای مختلف شامل کرنل‌ها و پارامترهای مختلف، در هر یک از ترکیب‌های مدل ماشین بردار پشتیبان مورد بررسی قرار گرفت. نتایج حاصل از بررسی بهترین ساختار، در ترکیب‌های مختلف مدل SVM برای چاه شماره یک در جدول (۴) آورده شده است. مطابق ترکیب شماره یک که فقط از داده‌های سطح آب زیرزمینی استفاده شده است، بهترین عملکرد مربوط به کرنل خطی با ضریب همبستگی ۰/۸۰۶ و جذر میانگین مربعات خطای ۰/۰۵۴۱ می باشد. در ترکیب ۲ و۳ که به ترتیب از داده های بارش و دما به صورت مجزا استفاده شده است، عملکرد مدل با توجه به ضریب همبستگی و جذر میانگین مربعات بدست آمده ضعیف بوده است. در ترکیب ۴ هم که از داده‌های تبخیر استفاده شد، عملکرد مدل ضعیف بوده و نشان‌دهنده این است که صرفاً استفاده از پارامترهای بارش، دما و یا تبخیر به هیچ وجه مناسب نبوده و در نتیجه باید از ترکیب‌هایی که به صورت ترکیبی از پارامترهای ذکر شده می‌باشد، استفاده نمود. در ترکیب ۵ که از داده‌های بارش، سطح آب زیرزمینی، دما و تبخیر استفاده شده است، عملکرد مدل نسبت به زمانی که از تک‌تک پارامترها استفاده شده بود بهتر شد. به همین دلیل در ترکیب ۶ و ۷ باز از همه پارامترها ولی با زمان تأخیرهای یک و دو ماهه استفاده شد. تا بتوان مقایسه بهتری برای دستیابی به ترکیب و ساختار برترانجام داد. در ترکیب ۶ از همه پارامترها با یک زمان تأخیر استفاده شد. در این ترکیب، ساختاری که از کرنل RBF استفاده شده بود با ضریب همبستگی ۰/۸۷۲ و جذر میانگین مربعات خطای شبیه‌سازی ۰/۰۵۵۰ بهترین عملکرد را داشت. در ترکیب ۷ از همه پارامترها با دو زمان تأخیر استفاده شد. در این ترکیب، ساختاری که از کرنل چند جمله‌ای با درجه چندجمله‌ای ۳، مقدار ۴۸۴/۳ استفاده شده بود با ضریب همبستگی ۰/۸۵۹ و جذر میانگین مربعات خطای شبیه‌سازی ۰/۰۴۱۴ بهترین عملکرد را داشت. بر اساس این جدول ترکیب ۶ دارای بهترین عملکرد برای مدل ماشین بردار پشتیبان بوده و بعد از آن ترکیب ۷ نتیجه بهتری دارد. به عبارت دیگر، افزایش تعداد تأخیرهای زمانی بیشتر از یک گام، باعث کاهش عملکرد مدل شده است. همچنین نتایج جدول (۴) نشان می‌دهد که استفاده از ساختارهای ترکیبی (یعنی ترکیب ۴، ۵، ۶ و ۷) دارای عملکرد بهتری نسبت به ترکیب‌هایی با صرفاً یک پارامتر به عنوان ورودی، می‌باشند. هم‌چنین نتایج نشان می دهد از بین پارامترهای ورودی، سطح آب زیرزمینی نسبت به دیگر پارامترها (بارش، دما و تبخیر)، دارای اهمیت بیشتری است. بهترین ساختار برای ترکیب‌های هفت‌گانه برای چاه‌های شماره ۲ تا ۵ در جدول (۵) ارائه شده است. با توجه به جدول (۵) برای چاه شماره ۲ ترکیب نوع ۶ که از داده های سطح آب زیرزمینی، بارش، دما و تبخیر با یک تأخیر زمانی استفاده شده، دارای بهترین عملکرد است. به عبارت دیگر داده سطح آب زیرزمینی، پارامتر اصلی و تأثیرگذار در خروجی می‌باشد. با توجه به جدول (۵)، برای چاه شماره ۳ ساختار برتر، ترکیب ۵ با کرنل خطی، ضریب همبستگی ۰/۸۷۷ و جذر میانگین مربعات خطای شبیه‌سازی ۰/۰۲۶۲ است. بر اساس جدول (۵) برای چاه شماره ۴ ترکیب ۶ دارای نتایج بهتری می‌باشد. برای چاه شماره ۵ نیز ترکیب ۶ دارای

نتایج بهتری بوده یعنی استفاده از پارامترهای دما، بارش و تبخیر باعث افزایش عملکرد مدل شده است. همان‌طوری که از بررسی معیارهای ارزیابی برای بهترین ساختار هر یک از ترکیب‌های هفت‌گانه برای چاه‌های مختلف مشخص شد کرنل (هسته)های RBF و Lin دارای بهترین عملکرد می‌باشند.

نتایج مدل‌سازی با مدل ماشین بردار پشتیبان- موجک

جدول (۶) بهترین عملکرد برای چاه شماره ۱ با تابع موجک Db4 به ازای درجه تجزیه ۲ با کرنل RBF را نشان می‌دهد. این ساختار دارای ضریب همبستگی ۰/۹۳۹و جذر میانگین مربعات خطای شبیه‌سازی ۰/۰۲۷۶ می‌باشد. با توجه جدول (۶) بهترین عملکرد برای چاه شماره ۲ با تابع موجک Db2 به ازای درجه تجزیه ۱ با کرنل RBF به دست آمد. این ساختار دارای ضریب همبستگی ۰/۹۸۰ و جذر میانگین مربعات خطای شبیه‌سازی ۰/۰۱۶۷ می باشد. بهترین عملکرد برای چاه شماره ۳ به ازای تابع موجک Db2 به ازای درجه تجزیه ۳ با کرنل RBF بدست آمد. این ساختار دارای ضریب همبستگی ۰/۹۵۰ و جذرمیانگین مربعات خطای شبیه‌سازی ۰/۰۱۷۶ می‌باشد. بهترین عملکرد برای چاه شماره ۴ با تابع موجک Coif1 به ازای درجه تجزیه ۳ با کرنل خطی بدست آمد. این ساختار دارای ضریب همبستگی ۰/۹۴۴و جذر میانگین مربعات خطای شبیه‌سازی ۰/۰۲۱۲ می‌باشد. با توجه به جدول (۶) بهترین عملکرد برای چاه شماره ۵ تابع موجک Coif1 به ازای درجه تجزیه ۲ با کرنل RBF بدست آمد. این ساختار دارای ضریب همبستگی ۰/۹۵۵ و جذر میانگین مربعات خطای شبیه سازی ۰/۰۰۷۹ می‌باشد. با توجه به جدول (۶) در تمامی سطوح تجزیه، برای چاه ۱ موجک مادر Db4، برای چاه ۴ و ۵ موجک مادر Coif1 و برای چاه ۲ و ۳ موجک مادر Db2 عملکرد بهتری داشته و به عنوان موجک‌های برتر شناخته می شوند.

مقایسه مدل‌های مختلف

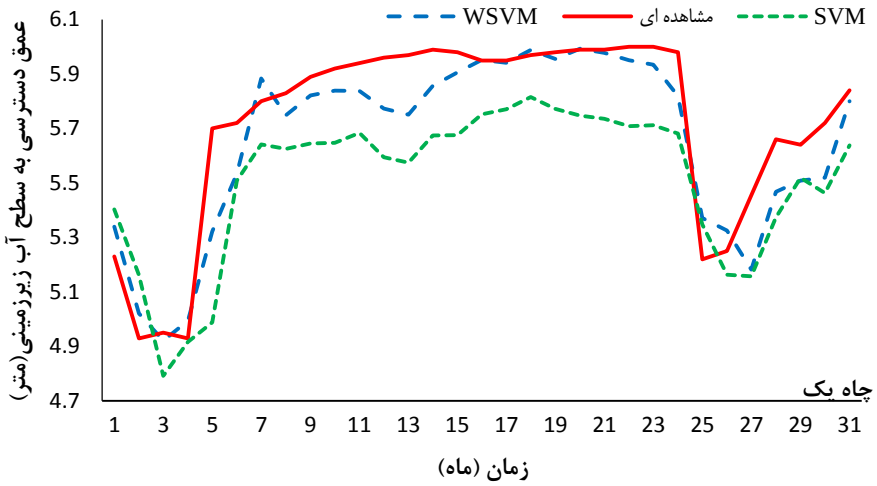
در جدول (۷) برای هر یک از چاه‌ها، دو مدل ماشین بردار پشتیبان و مدل ترکیبی ماشین بردار پشتیبان-موجک با هم مقایسه شده‌اند. مدل ترکیبی ماشین بردار پشتیبان-موجک نسبت به مدل ماشین بردار پشتیبان عملکرد بهتری داشته و این به دلیل آن است که از قبل، روی داده‌ها پردازش صورت می‌گیرد، به طوری‌که داده‌ها دسته‌بندی شده و با فرکانس های مختلف وارد ماشین بردار پشتیبان می شوند.

جدول ٦- نتایج WSVM برای موجک‌های مختلف برای چاه‌های مختلف.

چاه	موجک	درجه تجزیه	پارامترهای مدل			ضریب همبستگی		جذر میانگین مربعات خطا	
			کرنل	Gama	Sin2	درجه	آموزش	شبیه‌سازی	آموزش	شبیه‌سازی
۱	Db4	۲	RBF	۳۱۶/۷	۲۵۳/۷۷	----	۰/۹۸۰	۰/۹۳۹	۰/۰۱۸۲	۰/۰۲۷۶
۲	Db2	۱	RBF	۹۶۸۵	۱۳۴۱	----	۰/۹۷۶	۰/۹۸۰	۰/۰۱۷۲	۰/۰۱۶۷
۳	Db2	۳	RBF	۱۰۵۲	۱۷۸۳	----	۰/۹۸۵	۰/۹۵۰	۰/۰۱۵۴	۰/۰۱۷۶
۴	Coif1	۳	Lin	۰/۴۶۹۷	----	----	۰/۹۷۰	۰/۹۴۴	۰/۰۱۸۳	۰/۰۲۱۲
۵	Coif1	۲	RBF	۱۵۰۲۰	۱۲۰۲۴	----	۰/۹۸۶	۰/۹۵۵	۰/۰۱۹۰	۰/۰۰۷۹

جدول ۷-مقایسه نتایج حاصل از مدل‌های مختلف.

مدل	نام چاه	ضریب آکائیک		ضریب همبستگی		جذر میانگین مربعات خطا	
		آموزش	شبیه‌سازی	آموزش	شبیه‌سازی	آموزش	شبیه‌سازی
SVM	۱	۱۶۵/۳۴	۴۰/۷۹	۰/۹۱۹	۰/۸۷۲	۰/۰۳۵۷	۰/۰۵۵۰
WSVM	۱	۱۴۳/۹۵	۲۰/۹۳	۰/۹۸۰	۰/۹۳۹	۰/۰۱۸۲	۰/۰۲۷۶
SVM	۲	۱۶۰/۱۵	۳۲/۰۵	۰/۹۷۲	۰/۹۵۶	۰/۰۱۸۷	۰/۰۱۷۳
WSVM	۲	۱۵۹/۴۹	۳۱/۲۸	۰/۹۷۶	۰/۹۸۰	۰/۰۱۷۲	۰/۰۱۶۷
SVM	۳	۱۷۷/۳۲	۴۹/۶۷	۰/۹۵۸	۰/۸۷۹	۰/۰۲۵۵	۰/۰۲۷۸
WSVM	۳	۱۲۳/۰۳	۱/۶۰	۰/۹۸۵	۰/۹۵۰	۰/۰۱۵۴	۰/۰۱۷۶
SVM	۴	۱۶۴/۵۰	۳۹/۳۷	۰/۹۰۴	۰/۸۵۷	۰/۰۳۲۱	۰/۰۴۶۰
WSVM	۴	۱۲۸/۰۰	۲/۳۱	۰/۹۷۰	۰/۹۴۴	۰/۰۱۸۳	۰/۰۲۱۲
SVM	۵	۱۶۴/۱۴	۲۹/۳۳	۰/۹۶۳	۰/۹۱۴	۰/۰۳۰۷	۰/۰۱۳۱
WSVM	۵	۱۴۴/۴۷	۵/۸۴	۰/۹۸۶	۰/۹۵۵	۰/۰۱۹۰	۰/۰۰۷۹

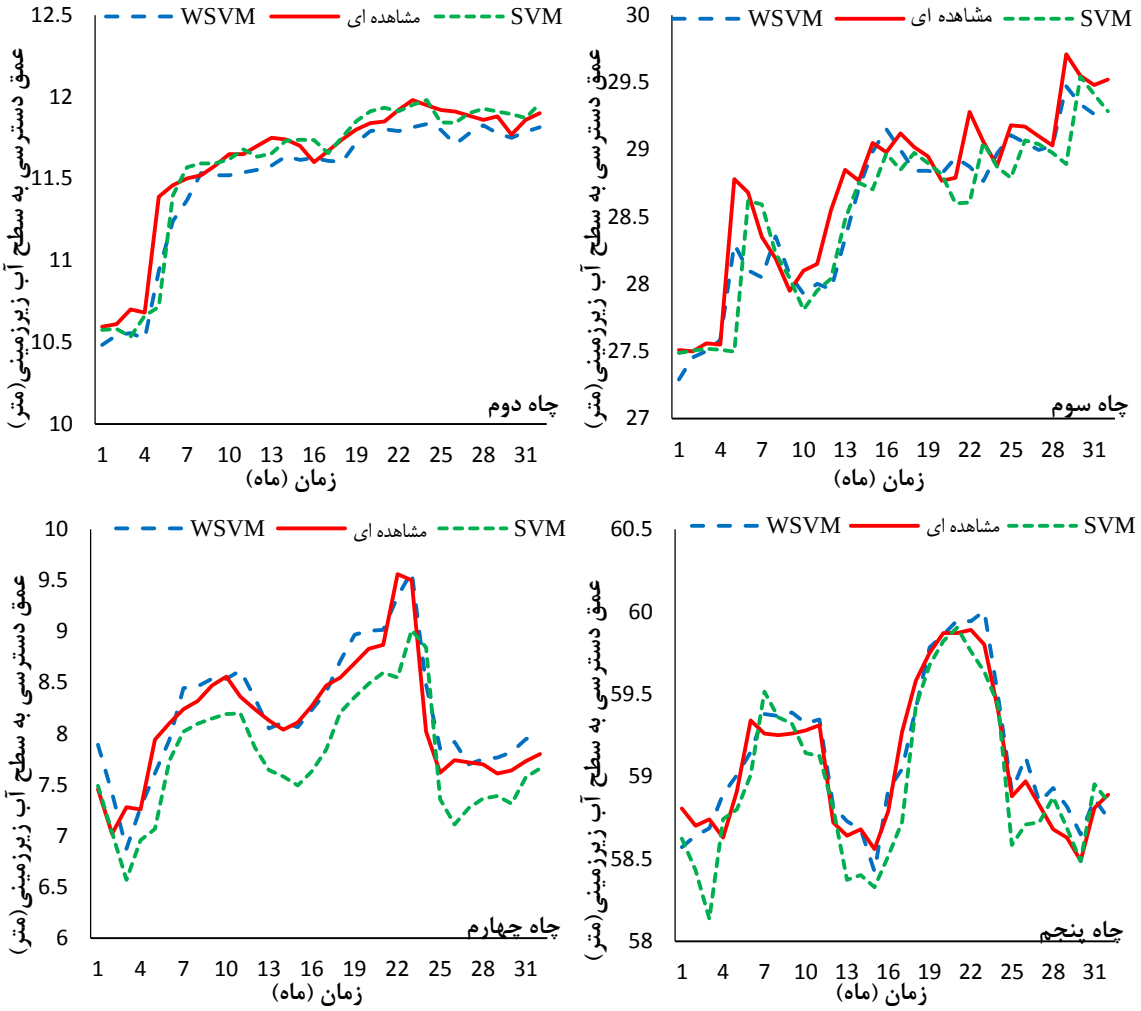


شکل۳- مقایسه نتایج حاصل از مدل‌های مختلف-مرحله شبیه‌سازی- چاه یک

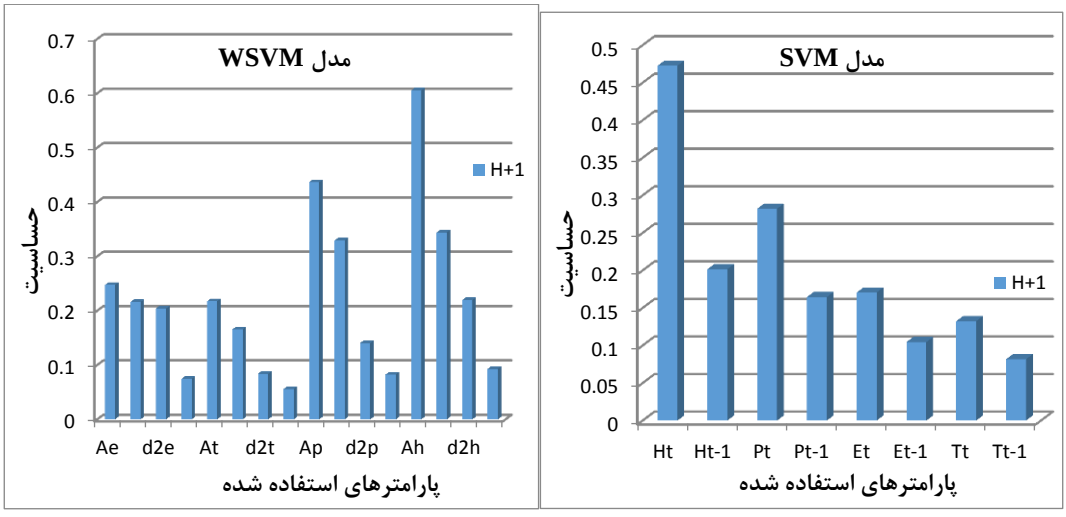
با توجه به جدول (۷) در تمامی چاه‌های مشاهده‌ای عملکرد مدل ترکیبی نسبت به مدل ماشین بردار پشتیبان بهتر بوده است. به‌طوری‌که در چاه شماره یک استفاده از مدل ترکیبی باعث بهبود عملکرد به میزان ۷ درصد، در چاه شماره ۲ به میزان ۳ درصد، در چاه شماره ۳ به میزان ۷ درصد، در چاه شماره ۴ به میزان ۸ درصد و در چاه شماره ۵ به میزان ۴ درصد افزایش ضریب همبستگی مشاهده می‌شود. همچنین با توجه به پارامتر جذر میانگین مربعات خطا نیز مشاهده می‌شود که مدل ترکیبی در تمامی چاه‌ها، دارای جذر میانگین مربعات خطای شبیه‌سازی کمتر بوده و در نتیجه عملکرد بهتری داشته است. بر اساس معیار آکائیک، مشاهده می‌شود که مدل ترکیبی در تمامی چاه‌ها دارای معیار آکائیک کمتر بوده و به عنوان مدل برتر شناخته می‌شود. شکل (۳) مقایسه مدل‌های استفاده شده در این تحقیق را برای مرحله شبیه‌سازی چاه شماره ۱ نشان می‌دهد. همان‌طوری‌که در این شکل مشخص است مدل ترکیبی در برآورد مقادیر حداکثر و حداقل سطح آب زیرزمینی نسبت به مدل ماشین بردار پشتیبان عملکرد بهتری داشته است. همچنین مشاهده می‌شود که مقادیر برآوردی مدل ترکیبی به مقادیر مشاهده‌ای نزدیک بوده و این مطلب را می‌رساند که از مدل ترکیبی، می‌توان برای پیش‌بینی مقادیر آتی سطح آب زیرزمینی بهره برد. در شکل (۴) مقایسه مدل‌های مختلف با مقادیر مشاهده‌ای برای چاه‌های شماره ۲ تا ۵ ارائه شده است.

بر اساس شکل (۴) با توجه به نزدیکی مقادیر دو مدل به مقادیر مشاهده‌ای برای چاه شماره ۲ عملکرد مدل ترکیبی بهتر می‌باشد. با توجه به شکل (۴) علیرغم نوسان در مقادیر برآوردی توسط مدل‌های مختلف، عملکرد مدل ماشین بردار پشتیبان از مدل ترکیبی برای چاه شماره ۳ بهتر بوده ولی در مجموع مقادیر مدل ترکیبی، به مقادیر مشاهده‌ای نزدیک‌تر بوده و باعث بهتر شدن عملکرد این مدل در مقابل مدل ماشین بردار پشتیبان شده است. برای چاه شماره ۴ مشخص است که مقادیر برآوردی مدل ترکیبی به مقادیر مشاهده‌ای نزدیک‌تر است. با توجه به این شکل به طور خیلی واضح مشخص است که عملکرد مدل ترکیبی بسیار بهتر از مدل ماشین بردار پشتیبان بوده است. برای چاه شماره ۵ با توجه به شکل (۴) مشاهده می شود که در برآورد نقاط ماکزیمم و مینیمم مقادیر مدل ترکیبی ماشین بردار پشتیبان-موجک به مقادیر مشاهداتی نزدیک‌تر بوده و بنابراین عملکرد بهتری داشته است. همچنین با بررسی ساختارهای مختلف برای تمامی چاه‌ها علیرغم اینکه تمامی کرنل‌ها دارای عمکرد مناسبی بودند ولی کرنلRBF و کرنل خطی در ساختارهای برتر نقش پررنگ‌تری را داشتند.

شکل (۵) تحلیل حساسیت را برای دو مدل ماشین بردار پشتیبان و مدل ترکیبی برای چاه شماره ۱ نشان می‌دهد. با توجه به شکل (۵) بیشترین تأثیر در سطح آب زیرزمینی مربوط به پارامتر سطح آب زیرزمینی و سپس بارش، تبخیر و دما می‌باشد. به عبارتی دما کمترین اثر و سطح آب زیرزمینی بیشترین تأثیر را دارند. بر اساس شکل (۵) نیز دیده می‌شود که پارامتر تقریب (Approximate) سطح آب زیرزمینی بیشترین تأثیر و به ترتیب پارامترهای بارش، تبخیر و دما قرار دارند. از بین پارامترهای جزئیات (Details) سطوح مختلف، سطوح ۱ دارای بیشترین اهمیت و سپس سطوح ۲ و سطوح ۳ کمترین تأثیر دارند.



شکل ٤- مقایسه نتایج حاصل از مدل‌های مختلف-مرحله شبیه‌سازی- چاه‌های ۲-۵.



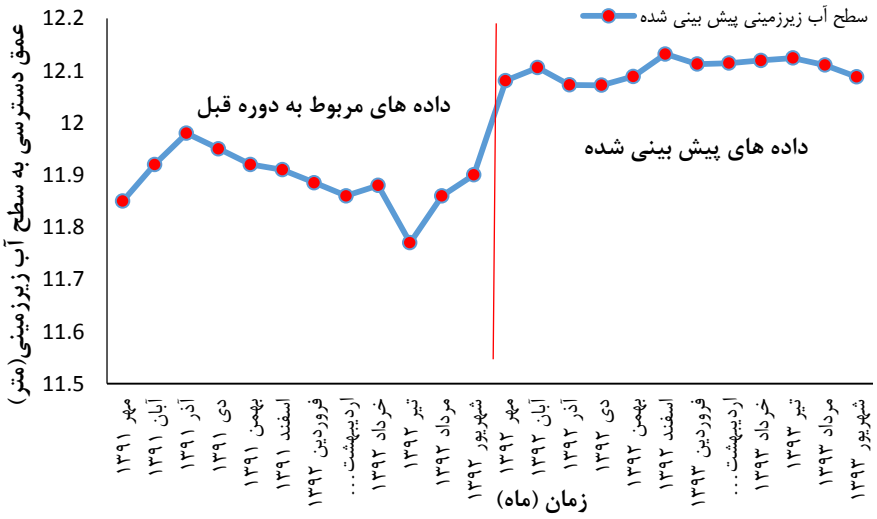
شکل ٥- تحلیل حساسیت مدل ماشین بردار پشتیبان و مدل ترکیبی

پیش‌بینی سطح آب زیرزمینی با مدل برتر

بعد از بررسی مدل‌های مختلف مشخص شد که مدل ترکیبی ماشین بردار پشتیبان- موجک، مدل برتر بوده به همین دلیل با تعیین بهینه پارامترهای مدل ترکیبی ماشین بردار پشتیبان- موجک، اقدام به پیش‌بینی سطح آب زیرزمینی شد. نتایج پیش‌بینی سطح آب زیرزمینی برای چاه شماره ۲ به عنوان نمونه از مهر ۱۳۹۲ تا شهریور ۱۳۹۳ در شکل (۶) ارائه شده است. برای بررسی روند تغییر سطح آب زیرزمینی در چاه شماره ۲ در شکل (۶) مقادیر سطح آب زیرزمینی از مهر ۱۳۹۱ به بعد آورده شده است. با بررسی روند سطح آب زیرزمینی در چاه شماره ۲ مشاهده می‌شود که عمق دسترسی به سطح آب در حال افزایش می‌باشد. با توجه به داده‌های پیش‌بینی شده، کاهش نیم‌متری آب زیرزمینی در طول یک سال، بیانگر بحران شدید در آینده برای آب زیرزمینی می‌باشد.

نتیجه‌گیری

در این تحقیق از مدل‌های ماشین بردار پشتیبان و ماشین بردار پشتیبان- موجک برای پیش‌بینی سطح آب زیرزمینی در ۵ حلقه چاه مشاهده‌ای در منطقه دشتستان استان بوشهر در دوره زمانی ماهانه استفاده شد. با بررسی و استفاده از تمامی پارامترهای مؤثر در مدل‌سازی ماشین بردار پشتیبان و مدل ترکیبی از توابع موجک گرفته تا کرنل‌های ماشین بردار پشتیبان، سعی شد مدل‌سازی در بهترین وضعیت صورت گرفته تا بتوان مقایسه‌ها در بهترین حالت ممکن انجام شوند. ترکیب مدل ماشین بردار پشتیبان با تبدیل موجک، یک ایده نو برای به وجود آمدن شبکه ای توانمندتر می‌باشد. پردازش داده‌های سری زمانی با آنالیز موجک باعث می‌شود که نقش و میزان تأثیر هر پارامتر در خروجی ماشین بردار پشتیبان به طور واضح بررسی گردد. در این تحقیق با توجه به آزمایش موجک‌های مختلف و توجه به نکته بالا، نتایج حاکی از آن است که موجک‌های Db2 و Coif1 دارای جواب‌های مناسبی می‌باشند. ضریب تعیین برای داده‌های آموزشی و شبیه‌سازی در مدل ماشین بردار پشتیبان- موجکی نسبت به مدل ماشین بردار پشتیبان بیشتر بوده که این موضوع بیانگر دقت روش ترکیبی می‌باشد. توانایی ماشین بردار پشتیبان- موجک در پیش‌بینی نقاط پیک و نوک تیز نمودار سری زمانی نسبت روش ماشین بردار پشتیبان بیشتر می‌باشد. با توجه به اینکه در مطالعات هیدرولوژی داشتن چنین نقاطی، ضروری است، اهمیت و نقش تبدیل موجک بیشتر نمایان می گردد. مقایسه نتایج حاصل از مدل‌سازی ماهانه حاکی از این است که استفاده از تبدیلات موجک در حالت ماهانه در بدترین شرایط، باعث افزایش ۳ درصدی عملکرد و در بهترین حالت باعث افزایش ۸ درصد عملکرد شده است. بررسی ساختارهای مختلف و تحلیل حساسیت نشان داد که از بین پارامترهای استفاده شده در این تحقیق پارامتر سطح آب زیرزمینی دارای بیشترین اهمیت و سپس پارامترهای سطح آب بارش، تبخیر و دما در مرحله بعدی اهمیت قرار دارند. با بررسی ترکیب‌های مختلف مشخص شد که استفاده از دو پارامتر دما و تبخیر علاوه بر پارامترهای اصلی سطح آب زیرزمینی و بارش باعث بهتر شدن نتایج شده است. با توجه به استفاده از کرنل‌های مختلف برای بدست آوردن نتایج بهتر، مشخص شد که دو نوع RBF و خطی دارای بهترین عملکرد بوده‌اند. با توجه به نتایج، عمق دسترسی به آب زیرزمینی در حال افزایش می‌باشد که این امر با توجه به بحران آب در کشور، باید بیش از پیش مورد توجه مسئولین استان بوشهر قرار گیرد.



شکل ٦- نتایج حاصل از پیش‌بینی و بررسی سطح آب زیرزمینی چاه ۲ (از مهر ۱۳۹۱ تا شهریور ۱۳۹۳).

منابع

1- Adamowski, J. and Chan, H.F., 2011. A wavelet neural network conjunction model for groundwater level forecasting. *Journal of Hydrology, 407(1–4)*, pp. 28-40.

2- Bierkens, M., 1998. Modeling water table fluctuations by means of a stochastic differential equation. *Water Resources Reserch, 34(10)*, pp. 2485-2499.

3- Cortes, C. and Vapnik, V., 1995. Support-Vector Networks. *Machine Learning, 20*, pp. 273-295.

4- Fallah-Mehdipour, E., Bozorg Haddad, O. and Mariño, M.A., 2013. Prediction and simulation of monthly groundwater levels by genetic programming. *Journal of Hydro-environment Research, 7(4)*, pp. 253-260.

5- Izadi, A.A., Davari, K., Alizadeh, A., Ghahreman, B. and Haghayeghi moghadam, S.A., 2007. Using Artificial Neural Network to Predict the Water Ttable (case study: Neyshabur plain). *Iranian journal of Irrigation and Drainage, 1(2)*, pp. 59-71. (In Persian).

6- Kumar, P.E. and Fofola-Georgiou, G., 1994. Wavelet in geophysiscs. New York, Academic Press.

7- Mallat, S.G., 1998. A wavelet tour of signal processing, *San Diego*, pp. 1-557.

8- Moosavi, V., Vafakhah, M., Shirmohammadi, B. and Ranjbar, M., 2014. Optimization of Wavelet-ANFIS and Wavelet-ANN Hybrid Models by Taguchi Method for Groundwater Level Forecasting. *Arabian Journal for Science and Engineering, 39*, pp. 1785–1796.

9- Nakhaei, M. Saberi nasr, R. and Faraj zadeh, R., 2011. Benefit of Wavelet-Artificial Neural Network in prediction of water table fluctuations. 4th Iran Water Resources Management Conference. *Amirkabir University of Technology, Tehran*. (In Persian).

10- Nayak, P., Satyaji Rao, Y.R. and Sudheer, K.P., 2006. Groundwater Level Forecasting in a Shallow Aquifer Using Artificial Neural Network Approach. *Water Resources Management, 20*, pp.77-90.

11- Nourani, V., Asghari Mogaddam, A. and Naderi, A.O., 2008. An ANN-based model for spatiotemporal groundwater level forecasting. *Hydrological Processes, 22*, pp. 5054–5066.

12- Nourani, V., Komasi, M. and Mano, A., 2009. A Multivariate ANN-Wavelet Approach for Rainfall–Runoff Modeling. *Water Resources Management, 23*, pp. 2877–2894.

13- Rajaee, T. and Zeynivand, A., 2015. Modeling of Groundwater Level using ANN–Wavelet Hybrid Model (Case Study: Sharif Abad Plain). *Journal of Civil and Environmental Engineering, 44(4)*, pp.51-63. (In Persian).

14- Riad, S., Mania, J., Bouchaou, L. and Najjar, Y., 2004. Rainfall-runoff model usingan artificial neural network approach. *Mathematical and Computer Modelling, 40(7–8)*, pp.839-846.

15- Schölkopf, B. and Smola, A.J., 2002. Learning with Kernels. *MIT Press*, Cambridge, MA.

16- Schölkopf, B., Smola, A. and Müller, K.R., 2005. Kernel principal component analysis. *Lecture Notes in Computer Science, 1327*, pp. 583-588.

17- Shiri, J., Kisi, O., Yoon, H., Lee, K.K. and Hossein Nazemi, A., 2013. Predicting groundwater level fluctuations with meteorological effect implications-A comparative study among soft computing techniques. *Computers & Geosciences, 56*, pp. 32-44.

18- Smola, A.J. and Schölkopf, B., 2004. Atutorial on support vector regression. *Statistics and Computing, 14(199)*, pp.199-222.

19- Solgi, A., 2014. Stream flow forecasting using combined Neural Network Wavelet model and comparsion with Adaptive Neuro Fuzzy Inference System and Artificial Neural Network methods (Case Study Gamasyab River,Nahavand). MSc. *Shahid Chamran University of Ahvaz, Iran*, pp.1-164. (In Persian).

20- Sujay Raghavendra, N. and Paresh Chandra, D., 2014. Support vector machine applications in the field of hydrology: A review. *Applied Soft Computing, 19*, pp.372-386.

21- Suryanarayana, C., Sudheer, C., Mahammood, V. and Panigrahi, B.K., 2014. An integrated wavelet-support vector machine for groundwater level prediction in Visakhapatnam, *India. Neurocomputing, 145*, pp.324-335.

22- Wang, W. and Ding, J., 2003. Wavelet Network Model and Its Application to the Prediction of Hydrology. *Natureand Science, 1(1)*, pp. 67-71.

Irrigation Sciences and Engineering Vol. 41, No. 1, Spring 2018, p. 165-180

EXTENDED ABSTRACT

Simulating Fluctuations of Groundwater Level Using a Combination of Support Vector Machine and Wavelet Transform

A. Eskandari [1], A. Solgi[2*] and Heidar Zarei[3]

1- Instructor, Department of Civil Engineering, Islamic Azad University of Bushehr, Iran.
2- * Corresponding author, Ph.D. Student, Faculty of Water Sciences Engineering, Shahid Chamran University of Ahvaz, Iran (*a-solgi@phdstu.scu.ac.ir*).
3- Associate Professor, Faculty of Water Sciences Engineering, Shahid Chamran University of Ahvaz, Iran.

. Received:2 January 2016 Accepted:10 December 2016

Keywords: Support vector machine, Wavelet transform, Groundwater tevel, Simulation.

Introduction

Groundwater resources are known as one of the most important water resources. Proper knowledge and exploitation of groundwater, particularly in arid and semi-arid regions, plays an important role in sustainable development in different fields such as agricultural, social, and economic activities. Many models have been used in order to predict groundwater level including empirical time series models and physical models (Bierkens, 1998). It is true that these models have been widely used, but when the dynamic behavior of a hydrological system is changed during the time, the mentioned models may not be able to predict water resource parameters, therefore making them unsuitable (Nayak et al. 2006). On the other hand, physical models need a great deal data to simulate groundwater fluctuations and, since the relations among effective parameters may be nonlinear, the mentioned models are not practical enough to show the existing relations.

Methodology

In recent years, nonlinear methods have been developed and the combination of different techniques to develop hybrid tools has improved the accuracy of the models which are used in different fields.

In the past years, scholarly research has been done in field of using ANN and ANFIS. Few research has, however, been done in the field of SVM. Therefore, in addition to SVM, the researchers in this study tried to use wavelet transform to predict groundwater levels in Dashtestan.

Since using direct data in any model can decrease the speed and accuracy of the model, standardization was used. Through this technique, each piece of data was converted into a number between 0 to 1 (Riad et al. 2004). Based on Solgi's (2014) study, the following equation was used to normalize the data.

$$y = 0.5 + (0.5 \times \left(\frac{x-\bar{x}}{x_{max}-x_{min}}\right)) \quad (1)$$

In this method, X is the original data, $\overline{X}$ is the average of the data, X_{max} is the maximum of the data, X_{min} is the minimum of the data, and y is the standardized data. In this study, 75% of the data was used for training and 25% for simulation.

SVM is one of the methods for supervised learning. SVM is a classifier which is known as a part of Kernel Methods for the learning process. These types of systems are used for classifying as well as estimating the fitting data functions in order to produce minimum errors. This method is a relatively new one which has been more practical compared with the older models like Perceptron neural networks. The approach in SVM is the linear classifying of data. In classifying, the researchers try to select a line which has the highest interval confidence. Finding the optimal line for data is done by the QO method which is a well-known method for solving problems with restrictions. Before linear classification, the data is shifted to a high dimensional space by phi function in order to solve data with a high complexity. In order to solve problems with huge dimensions, Lagrange Duality Theorems is used to convert a desirable minimization issue to its double form which uses a simple function, called core function. It presents Phi function multiply instead of using a complex function that shifts to high dimensional space.

With regard to wavelet transform, many books and articles have been published which can be used to understand the theory of wavelet transform. The wavelet function is a function which has the two features of oscillation and short-time period.

. Ψ(x(is a wavelet function if and only if its Fourier transform has the features presented in Equation 2.

$$C_g = \int_{-\infty}^{+\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty \quad (2)$$

This condition is known as the acceptance condition. The above equation can be equal to the following equation. In other words, Equation 6 has the wavelet function condition if Equation 3 is established.

$$\int_{-\infty}^{+\infty} \Psi(x) dx = 0 \quad (3)$$

This feature, whose average equals to zero, is not a very strong limitation. Based on this, many functions can be considered as the wavelet function. Ψ(x(is the mother wavelet function if the size and location of the used functions are changed regarding the two mathematical functions of Translation and Scale in the studied signals. Finally, the coefficients of the wavelet at every point of the signal (b) and every scale (a) can be calculated through Equation 4.

$$\Psi_{a,b}(x) = \frac{1}{\sqrt{a}} \Psi\left(\frac{x-b}{a}\right) \quad (4)$$

Results and Discussion

Regarding the best structure of each used model, WSVM showed a better performance compared with SVM because of data pre-processing; the data was classified and then entered into the model. This can be concluded based on R^2 in the two steps of training and testing. Analysis of different structures proved that the data of groundwater levels has the highest importance for modeling. The data regarding precipitation, evaporation, and temperature were found to be the less important data, respectively.

Conclusions

In this study, SVM and WSVM were used for predicting groundwater levels in five observation wells, located in Dastestan Plain, Bushehr, in a monthly scale. By detecting and using all effective parameters in SVM modeling and a hybrid SVM and WSVM, attempt was made to arrive at a model which gives the best results under the best conditions. Comparison of monthly simulations showed that using the wavelet transform increases the performance by 3%

to 8% in the best condition. By detecting different combinations, it was concluded that using temperature and evaporation as well as groundwater level and precipitation improves the results. Based on using different kernals to reach better results, it was seen that the RBF and the linear model have the best performance. Regarding the results, the depth of access to groundwater increases and, regarding the water crisis in Iran, decision makers should pay more attention to this issue in Bushehr province.

References

1- Bierkens, M., 1998. Modeling water table fluctuations by means of a stochastic differential equation. *Water Resources Reserch,34(10)*, pp.2485-2499.

2- Nayak, P., Satyaji Rao, Y.R. and Sudheer, K.P., 2006. Groundwater Level Forecasting in a Shallow Aquifer Using Artificial Neural Network Approach. *Water Resources Management*, *20*, pp.77-90.

3- Riad, S., Mania, J., Bouchaou, L. and Najjar, Y., 2004. Rainfall-runoff model usingan artificial neural network approach. *Mathematical and Computer Modelling*, *40(7–8)*, pp.839-846.

4- Solgi, A., 2014. Stream flow forecasting using combined Neural Network Wavelet model and comparsion with Adaptive Neuro Fuzzy Inference System and Artificial Neural Network methods (Case Study:Gamasyab River,Nahavand). MSc. *Shahid Chamran University of Ahvaz*, Iran (in Persian with English abstract), pp.1-164.