

مقدمه

شرایط اقلیمی خشک و نیمه‌خشک حاکم بر کشور ایران، همیشه به عنوان یک عامل اصلی در محدودیتهای منابع آب بوده‌است. از طرف دیگر کاهش میزان بارندگی به جهت خشکسالی در چندین سال اخیر، اعمال مدیریت صحیح بر منابع را ضرورت بخشیده است. از آنجا که منابع آب زیرزمینی بیشترین سهم را در تأمین آب دارا می‌باشند، شناخت صحیح و بهره‌برداری اصولی از آن‌ها می‌تواند در توسعه پایدار فعالیتهای اجتماعی و اقتصادی نقش بسزایی ایفا نماید [۱۰]. در مناطق خشک و نیمه‌خشک، آبهای زیرزمینی تنها منبع مورد اعتماد محسوب می‌گردند. با توجه به این که ریزش‌های جوی در زمان‌های محدود اتفاق می‌افتد، لذا جریانهای سطحی عموماً به صورت سیلابی و پس از هر بارش مشاهده و سپس قطع می‌گردد. بنابراین کلیه برنامه‌ریزیهای کوتاه‌مدت و بلندمدت تأمین آب در مناطق خشک و نیمه‌خشک نیز بر مبنای حجم آب زیرزمینی می‌باشد [۱۳]. مدل‌های رایج که در این زمینه توسعه یافته‌اند از جمله مدل‌های رگرسیونی با سری زمانی، مدل‌های فیزیکی و مدل‌های هوشمند را می‌توان نام برد. مدل‌های هوشمند به دلیل سرعت بالای پردازش و دقت خوب و هزینه کم نسبت به مدل‌های دیگر گسترش زیادی یافته‌اند که برای شبیه‌سازی پارامترهای منابع آب مناسب می‌باشند [۱۷]. پارامترهای بسیاری بر پدیده‌های هیدرولوژیکی مؤثرند و تعیین تأثیر میزان این پارامترها نامشخص است. در این صورت، تعیین بهترین ترکیب از پارامترهای ورودی به مدل هوشمند امری وقت‌گیر می‌باشد. کاربرد روش‌های جدید باعث کاهش سعی و خطا شده و به روند مدل‌سازی سرعت بخشیده است. در این پژوهش از آزمون گاما برای پیش‌پردازش داده‌ها استفاده گردید. آزمون گاما نخستین بار توسط کانسکر در سال ۱۹۹۷ به صورت خلاصه معرفی گردید و در سالهای بعد توسط دورانت، جونس و همکاران و ایوانس با جزئیات مورد بررسی قرار گرفت [۱۶]. این روش یک ابزار مدل‌سازی غیرمتغیری است که به کمک آن می‌توان ترکیب مناسب از پارامترهای ورودی را به منظور مدل‌سازی داده‌های خروجی و ایجاد یک مدل هموار قبل از ایجاد مدل بررسی نمود. به طوری که با بررسی مقدار میانگین مربعات خطا در هر ترکیب، یک معیار مناسب برای قضاوت در مورد آنها فراهم می‌گردد و در نهایت ترکیبی که دارای کمترین مقدار خطا باشد، به عنوان ترکیب بهینه مورد استفاده قرار می‌گیرد [۱۴]. بسالت

 کاربرد شبکه بیزین و مدل ماشین بردار پشتیبان
در پیش‌بینی تغییرات سطح تراز ایستابی
(مطالعه موردی: دشت اردبیل)

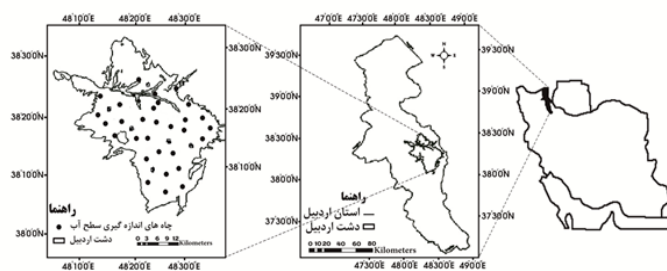
فاطمه آخونی پورحسینی^۱ و اسماعیل اسدی^۲
تاریخ دریافت: ۱۳۹۴/۱۱/۲۰ تاریخ پذیرش: ۱۳۹۵/۰۹/۰۱

چکیده:

آب‌های زیرزمینی به عنوان یکی از منابع مهم و عمده تأمین آب شرب و کشاورزی، به ویژه در مناطق خشک و نیمه‌خشک مطرح بوده است. شبیه‌سازی سیستم آب‌های زیرزمینی به دلیل پیچیدگی این سیستم‌ها به آسانی میسر نیست. در این مقاله با استفاده از داده‌های سطح تراز ایستابی دشت اردبیل در بازه زمانی (۱۳۹۰-۱۳۵۱)، به ارزیابی عملکرد آزمون گاما برای پردازش و انتخاب ورودی‌های مناسب و کارایی مدل‌های حداقل مربعات ماشین بردار پشتیبان و شبکه بیزین پرداخته شد. پارامترهای سطح تراز ایستابی ماهانه با تأخیرهای مختلف به عنوان ورودی آزمون گاما در نظر گرفته شد. نتایج آزمون گاما نشان داد که سطح تراز ایستابی با شش تأخیر زمانی، نتایج بهتری به منظور پیش‌بینی ارائه می‌دهد. شبیه‌سازی سطح تراز ایستابی با استفاده از دو مدل حداقل مربعات ماشین بردار پشتیبان و شبکه بیزین نیز نشان داد که بهترین ساختار ورودی برای پیش‌بینی سطح تراز ایستابی ماه بعد، تا شش تأخیر زمانی خواهد بود. از میان دو مدل با ساختار ورودی یکسان، مدل حداقل مربعات ماشین بردار پشتیبان، عملکرد بهتری را با توجه به ضریب تبیین ۰/۹۷۷، میانگین خطای مطلق ۰/۲۰۴ و جذر میانگین مربعات خطای ۰/۳۰۷، نسبت به شبکه بیزین داشته است. نتایج تحقیق نشان داد که آزمون گاما در انتخاب ترکیب ورودی مناسب در محاسبات نرم می‌تواند کارایی بهتری داشته باشد.

کلید واژه‌ها: آزمون گاما، سطح تراز ایستابی، دشت اردبیل، شبکه بیزین، ماشین بردار پشتیبان

۱. نویسنده مسئول، دانشجوی ارشد مهندسی منابع آب، دانشکده کشاورزی، دانشگاه تبریز، پست الکترونیکی fateme.pourhosseini@yahoo.com
۲. عضو هیئت علمی گروه مهندسی آب، دانشگاه تبریز، پست الکترونیکی: esasadi@gmail.com



شکل ۱- موقعیت جغرافیایی منطقه مورد مطالعه

Figure 1. The geographical location of the study area

یک مدل پیش‌بینی بر اساس خشکسالیهای گذشته برای پیش‌بینی خشکسالی در آینده از شبکه بیزین استفاده کردند. در این تحقیق یک مدل پیش‌بینی آماری چندمتغیره با توجه به میزان رواناب در حوضه طرح‌ریزی شد که مهم‌ترین مزیت مدل داشتن ویژگی‌های احتمالی در تجزیه و تحلیل خشکسالی در سالهای آتی بود. در این مدل بالاترین پیش‌بینی احتمالی با یک ارزیابی از عدم قطعیتها، مورد تحلیل قرار گرفت. شبکه‌های بیزین احتمال خشکسالی در سالهای آینده را با شدت‌های مختلف نشان دادند. فرمانی و همکاران [۷] در منطقه کپنهاگ از شبکه‌های بیزین برای مدیریت بهینه آلودگی آبهای زیرزمینی استفاده کردند. نتایج به‌دست‌آمده حاکی از کارایی مناسب الگوهای بیزین در مدیریت آلودگی آبهای زیرزمینی منطقه بوده است.

هدف از تحقیق حاضر، ارزیابی عملکرد آزمون گاما برای پردازش و انتخاب ورودیهای مناسب و کارایی مدل حداقل مربعات ماشین بردار پشتیبان و شبکه بیزین در پیش‌بینی سطح تراز ایستابی دشت اردبیل است.

مواد و روش

منطقه مورد مطالعه و داده‌های مورد استفاده

دشت اردبیل در شمال غربی ایران واقع شده است. این دشت در عرض جغرافیایی ۳۸ درجه و ۳ دقیقه تا ۳۸ درجه و ۲۷ دقیقه شمالی و طول جغرافیایی ۴۷ درجه و ۵۵ دقیقه تا ۴۸ درجه و ۲۰ دقیقه شرقی واقع شده است. مساحت آن حدود ۹۹۰ کیلومتر مربع می‌باشد. در محدوده دشت اردبیل تعداد ۲۲۴۳ حلقه چاه عمیق و نیمه عمیق، ۲۰ دهنه چشمه و ۱۸ رشته قنات وجود دارد که به مصارف کشاورزی اختصاص دارند. مجموع برداشت سالانه از چاه‌ها به منظور آبیاری دشت در حدود ۱۷۷ میلیون مترمکعب است همچنین برداشت از چشمه‌ها و قنات به منظور آبیاری دشت به ترتیب در حدود ۱/۲ و ۰/۹ میلیون مترمکعب در سال است که در شکل ۱ موقعیت جغرافیایی منطقه مورد مطالعه قابل مشاهده می‌باشد.

برای اجرای هر تحقیق لازم است روشی برای رسیدن به اهداف تحقیق طراحی شود. هر اندازه روش طراحی شده متناسب با ویژگیهای ذاتی و قابل حصول داده‌ها باشد، موجب رسیدن به

پور و همکاران [۱] در انتخاب بهینه پارامترهای ورودی به منظور مدل‌سازی مقاومت برشی خاک، از این آزمون استفاده نمودند که نتایج نشان دهنده کارایی، دقت بالای این روش در انتخاب، کاهش حجم محاسبات و صرفه‌جویی در وقت و هزینه‌ها می‌باشد. قبائی سوق و همکاران [۸] آزمون گاما را در پیش‌پردازش میان ۳۱ ترکیب مختلف پارامترهای ورودی جهت مدل‌سازی تابش خورشیدی مورد استفاده قرار دادند، همچنین با استفاده از آزمون M تعداد داده مورد نیاز برای آموزش مدل نیز تعیین گردید و در ادامه این روش با روش رگرسیون محلی و شبکه عصبی مصنوعی^۱ مقایسه شد که نتایج، کاربردی بودن این روش را نشان داد. سیفی و ریاحی مدوار [۱۹] به منظور پیش‌بینی تخریب-تورق پتانسیل از پیش‌پردازش آزمون گاما در مدل‌های حداقل مربعات ماشین بردار پشتیبان^۲ و سیستم استنتاج تطبیقی عصبی-فازی^۳ و شبکه عصبی مصنوعی استفاده نمودند که نتایج نشان‌دهنده دقت بالای LSSVM-GT و کاهش مراحل سعی و خطا در مدل‌سازی بود. ملیت و همکاران [۱۲] در تحقیقی با استفاده از مدل حداقل مربعات پشتیبان، به پیش‌بینی کوتاه‌مدت سری زمانی داده‌های هواشناسی نظیر تابش خورشیدی، دمای هوا، رطوبت نسبی پرداختند. نتایج به دست آمده از تحقیق آنها نشان داد که در میان روشهای مختلف شبکه‌های عصبی مصنوعی (شبکه‌ی معمولی، پرسپترون چندلایه، توابع پایه شعاعی^۴ و شبکه عصبی احتمالی) و مدل حداقل مربعات پشتیبان، مدل حداقل مربعات پشتیبان به‌عنوان مدل برتر معرفی گردید. بیوندی و همکاران [۲] در تحقیقی در جنوب ایتالیا با استفاده از شبکه‌های بیزین اقدام به ارزیابی عدم قطعیت موجود در پیش‌بینی زمان واقعی رخداد سیلاب پرداختند که نتایج آن‌ها حاکی از ترکیب عدم قطعیت در پارامترهای توزیع بارندگی و سیلاب مشاهداتی می‌باشد. همچنین نتایج آنها نشان داد که مدل‌های بیزین، توانایی بالایی در پیش‌بینی رخداد سیلاب و زمان آن دارند. مددگر و همکاران [۱۵] در بررسی تغییرات مکانی خشکسالی، در رودخانه گانیسن^۵ ایالات متحده آمریکا و توسعه

- 1- Artificial Neural Network
- 2- Least Square Support Vector Machine
- 3- Adaptive Network Fuzzy Inference System
- 4- Radial Basis Function
- 5- Gunnison

Table 1. Cross-correlation coefficient between input and output parameters

$G_{(t-6)}$	$G_{(t-5)}$	$G_{(t-4)}$	$G_{(t-3)}$	$G_{(t-2)}$	$G_{(t-1)}$	G_t
0.42	0.55	0.63	0.74	0.85	0.97	

گاما بر مبنای مجموعه $N[i, k]$ می‌باشد. به عنوان k امین نزدیک‌ترین همسایگی‌ها برای بردار مشاهداتی ورودی تعریف شد به طوری که می‌باشند. p حداکثر تعداد همسایگی‌ها است که معمولاً بین ۱۰ تا ۵۰ در نظر گرفته می‌شود [۴]. به منظور تعیین مقدار آماره گاما (Γ)، ابتدا مقدار، با توجه به داده‌های ورودی به شرح زیر محاسبه می‌شود [۲۰]:

$$\delta_M(k) = \frac{1}{M} \sum_{i=1}^M |x_{N[i, k]} - x_i|^2, 1 \leq k \leq p \quad (3)$$

که در آن نماد $|\dots|$ به عنوان فاصله اقلیدسی بردار $x_{N[i, k]}$ از بردار همسایگی‌های آن می‌باشد. همچنین مقدار آماره‌ی گامای ($\gamma_M(k)$) کوچک، با استفاده از داده‌های خروجی به شرح زیر به دست می‌آید:

$$\gamma_M(k) = \frac{1}{2M} \sum_{i=1}^M |y_{N[i, k]} - y_i|^2, 1 \leq k \leq p \quad (4)$$

که در آن $\gamma_{N[i, k]}$ مقدار خروجی نظیر k امین نزدیک‌ترین همسایگی برای بردار x_i می‌باشد. به این ترتیب p مقدار برای $\delta_M(k)$ و مقدار برای $\gamma_M(k)$ به دست آمد. سپس بین p زوج مرتب $\{\delta_M(k), \gamma_M(k)\}$ رابطه رگرسیونی خطی ساده به شرح رابطه ۵ برقرار می‌شود طوری که در آن، δ و γ به ترتیب به عنوان متغیرهای وابسته و مستقل در نظر گرفته می‌شوند:

$$\gamma = A \delta + \Gamma \quad (5)$$

مقدار آماره‌ی Γ در واقع عرض از مبدأ خط رگرسیون فوق می‌باشد. A شیب خط رگرسیون بوده که نشان‌دهنده پیچیدگی مدلی است که از روی داده‌های مشاهداتی ایجاد می‌شود [۶]. متغیر دیگری که معمولاً در آزمون گاما به کار برده می‌شود متغیر بدون بعد V_{ratio} نام دارد که مقدار آن بین صفر و یک بوده و از رابطه ۶ محاسبه می‌شود [۶]:

$$V_{ratio} = \frac{\Gamma}{\sigma^2(y)} \quad (6)$$

که در آن $\sigma^2(y)$ واریانس مقادیر خروجی مشاهداتی می‌باشد. هر چه مقدار V_{ratio} به صفر نزدیک‌تر باشد ترکیب انتخابی برای مدل‌سازی، ترکیب مناسب‌تری است. اگر این مقدار از عدد یک کم گردد مقدار ضریب تبیین مدل به دست خواهد آمد. اگر فرض شود که N تعداد متغیرهای ورودی مؤثر بر خروجی مدل باشد در این صورت، تعداد $2N - 1$ حالت (ترکیب‌های مختلفی از متغیرهای ورودی) وجود دارد که می‌توان از هر کدام برای تهیه مدل استفاده کرد. استفاده از همه ترکیبات ممکن برای یافتن بهترین

نتایج بهتر و دقیق‌تر نیز می‌شود. متغیرهای ورودی در نظر گرفته شده علاوه بر مؤثر بودن بر متغیرهای خروجی از نظر اندازه‌گیری و جمع‌آوری نیز در حد امکان آسان و قابل دسترسی هستند. به منظور تخمین سطح تراز ایستابی برای ماه‌های آینده، از اطلاعات سطح تراز ایستابی ماه‌های قبل در دشت اردبیل استفاده شده‌است. با استفاده از ماتریس همبستگی سری زمانی با تأخیر زمانی شش ماه قبل، به عنوان ورودی به مدل‌ها معرفی شد. مقدار تراز سطح ایستابی ماه بعد (G_t) به عنوان خروجی مدل‌ها تعیین گردید شش متغیر شامل سری مربوط به سطح ایستابی با تأخیر یک ماه ($G(t-1)$ ، دو ماه ($G(t-2)$ ، سه ماه ($G(t-3)$ ، چهار ماه ($G(t-4)$ ، پنج ماه ($G(t-5)$ و شش ماه ($G(t-6)$) به عنوان متغیر ورودی به مدل معرفی گردید (جدول ۱).

آزمون گاما

آزمون گاما نشان دهنده آن قسمت از واریانس خروجی است که نمیتوان با اعمال هر مدل هموار بر ورودیها محاسبه شود. با این آزمون مقدار میانگین مربعات خطای مدل پیش از استفاده محاسبه و ترکیب بهینه متغیرهای ورودی شناسایی می‌شود. بنابراین، آزمون گاما را می‌توان به عنوان ابزاری مناسب برای یافتن بهترین ترکیب از متغیرهای ورودی به مدل غیرخطی در نظر گرفت. در واقع آزمون گاما تخمینی از حداقل میانگین مربعات خطای آن بخش داده‌ها است که با مدل خطی نمی‌توان آن را تعیین کرد. فرض می‌شود مجموعه داده‌های مشاهداتی به صورت زیر باشد [۱۷]:

$$\{x_i - y_i, 1 \leq i \leq M\} \quad (1)$$

که در آن بردار مشاهداتی ورودی (مانند رواناب، بارش، درجه حرارت هوا با تأخیرهای مختلف و هر متغیر دیگر مؤثر)، خروجی مشاهداتی مانند میانگین دبی رواناب روزانه و M تعداد کل مشاهدات می‌باشد. از بردار x_i (در فضای m بعدی) برای مدل‌سازی y_i (در فضای R) استفاده می‌شود. بنابراین، با توجه به فرض فوق رابطه‌ی بین ورودی و خروجی به شرح زیر در نظر گرفته می‌شود:

$$y = f(x) + r \quad (2)$$

که در آن f یک تابع هموار و r متغیر تصادفی نشان‌دهنده خطا می‌باشد. همچنین فرض می‌شود میانگین توزیعی که r از آن پیروی می‌کند صفر و واریانس خطا، $\text{Var}(r)$ کرداندار است. به این ترتیب مدل نشان داده شده با رابطه ۲ به مدلی هموار با مشتقات جزئی مرتبه اول محدود می‌گردد.

فرض می‌شود برای بردار ورودی X_i ، مجموعه $N[i, k]$ ($1 \leq k \leq p$) به نام مجموعه نزدیک‌ترین همسایگی است. پایه و اساس آزمون

رابطه ریاضی Mathematical Equation	توابع کرنل Kernel Functions
$K(x_i, x_j) = x_i^T \cdot x_j$	تابع کرنل خطی Linear
$K(x_i, x_j) = (x_i^T \cdot x_j + \tau)^d$	تابع کرنل چندجمله‌ای Polynomial
$K(x, x_i) = \exp\left(\frac{- x_i - x_j ^2}{2\sigma^2}\right)$	تابع کرنل با پایه شعاعی Basis Radial Function

$$\frac{\partial L}{\partial W} = 0 \rightarrow W = \sum_{i=1}^N \alpha_i \phi_i(x_i)$$

$$\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^N \alpha_i = 0 \quad (11)$$

$$\frac{\partial L}{\partial e_i} = 0 \rightarrow \alpha_i = \gamma e_i, i = 1, \dots, N$$

$$\frac{\partial L}{\partial \alpha_i} = 0 \rightarrow y_i = W^T \cdot \phi(x_i) + b + e_i, \\ i = 1, \dots, N$$

با استفاده از شرایط KKT مدل حداقل مربعات ماشین بردار پشتیبان به منظور تابع تخمین به صورت رابطه ۱۲ نتیجه می‌گردد:

$$y_X = \sum_{i=1}^N \alpha_i K(X, X_j) + b \quad (12)$$

تابع کرنل نامیده می‌شود که با تبعیت از شرایط Mercer به عنوان تابعی با ایجاد ضریب داخلی در فضای ویژگی معرفی می‌شود [۱۸].

$$K(X_i, X_j) = \langle \phi(x_i), \phi(x_j) \rangle, i, j = 1, \dots, N \quad (13)$$

برای تعیین پارامترت تنظیم کننده γ و تعیین $\sigma < 0$ مربوط به توابع کرنل (نظیر RBF) نیاز است. انتخاب بهترین تابع کرنل به وسیله سعی و خطا امکان پذیر است. در این تحقیق از سه تابع مختلف استفاده شده است که در جدول ۲ آورده شده است:

شبکه بیزین^۳ (BN)

شبکه بیزین یک الگوی نموداری احتمالاتی است که مجموعه‌ای از متغیرها و احتمالات مربوط به هر یک را نشان می‌دهد. یک گراف مستقیم و چرخه‌ای است که در آن، گره‌ها در حکم متغیرهای مسئله هستند. ساختار یک شبکه بیزین در واقع یک نمایش نموداری از اثرات متقابل متغیرهایی است که باید الگوبندی شوند و علاوه بر اینکه کیفیت رابطه بین متغیرهای مسئله را نشان می‌دهد، کمیت ارتباط بین این متغیرها را نیز به نمایش می‌گذارد که به صورت عددی از توزیع احتمال مشترک آنها استفاده می‌کند. این روش بر مبنای محاسبات احتمالات شرطی (قانون بیز) بوده که معادله زیر رابطه بیز را نشان می‌دهد [۵].

$$P(a|b) = \frac{P(b|a) * P(a)}{P(b)} \quad (14)$$

که در آن $P(a)$ احتمال وقوع پیشامد a ، $P(b)$ احتمال وقوع پیشامد b ، $p(a|b)$ احتمال وقوع پیشامد a به شرطی که پیشامد b اتفاق بیافتد، $p(b|a)$ احتمال وقوع پیشامد b به شرطی که پیشامد a اتفاق بیافتد. هر شبکه بیزین از سه جزء اصلی شامل مجموعه‌ای از گره‌ها، مجموعه‌ای از بندها و مجموعه‌ای از احتمالات تشکیل شده است. شکل ۲ نمونه‌ای از ساختار شبکه بیزین را نشان می‌دهد.

ترکیب ورودی به مدل بسیار وقت گیر و خسته کننده است. بنابراین در شرایطی که تعداد متغیر ورودی زیاد باشد با استفاده از آزمون گاما می‌توان بهترین ترکیب از میان حالات ممکن را تشخیص داد [۱۱].

مدل حداقل مربعات ماشین بردار پشتیبان (LS-SVM)

مجموعه داده‌های $\{x_i, y_i\}_{i=1}^N$ داده شده است که $x_i \in R^p$ به عنوان بردار ورودی $y_i \in R$ به عنوان بردار خروجی معرفی می‌شود. تابع رگرسیون غیرخطی در مسائل تخمین به صورت رابطه ۷ می‌باشد [۱۷]:

$$f(x) = W^T \cdot \phi(x) + b \quad (7)$$

W و b به ترتیب مقادیر وزنها و بایاس تابع رگرسیون هستند. $\phi(x)$ ترسیم غیرخطی ورودی‌ها در فضای ویژگی با ابعاد بالا است. حل مسئله رگرسیون غیرخطی بر اساس معادله بهینه‌سازی در رابطه ۸ حل می‌گردد:

$$\min_{w, e, b, j} (W, e) = \frac{1}{2} W^T W + \frac{\gamma}{2} \sum_{i=1}^N e_i^2 \quad (8)$$

بر اساس محدودیت:

$$y_i = W^T \cdot \phi(x) + b + e_i \quad i = 1, \dots, N \quad (9)$$

γ پارامتر تنظیم کننده خطاست و میزان خطا را نشان می‌دهد. با استفاده از شکل لاگرانژی از تابع هدف اصلی می‌توان نوشت:

$$L(w, b, e, \alpha) = j(W, e) - \sum_{i=1}^N \alpha_i \{W^T \cdot \phi(x_i) + b + e_i - y_i\}$$

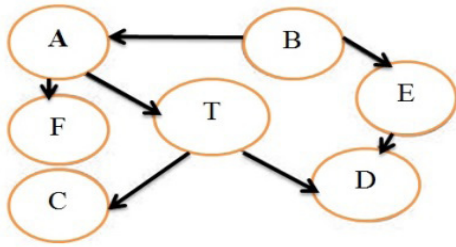
α_i ضریب لاگرانژ است بر اساس شرایط کان-تاکر^۱ (KKT):

1- Least Squares Support Vector Machine

2- Kuhn-Tucker

جدول ۳- مشخصات آماری سطح تراز ایستابی دشت اردبیل
Table 3. Statistical Profile water level plain of Ardabil

پارامتر	کل	آموزش (متر)	صحت سنجی (متر)	Parameter	Test(m)	Train(m)	Total
میانگین	Mean	1334.2	1336.8	1328.2			
حداقل	Min	1325.1	1330.9	1325.1			
حداکثر	Max	1339.3	1339.3	1332.03			



شکل ۲- نمایش گراف مستقیم و چرخه‌ای شبکه بیزین
Figure 2. Bayesian network graph display directly and cyclic

$$X_n = \frac{X_i - X_{max}}{X_{min} - X_{max}} \quad (15)$$

در رابطه ۱۵، X_i داده‌های مشاهداتی، X_{min} و X_{max} به ترتیب معرف داده‌های حداکثر و حداقل می‌باشد. بعد از تعیین مقادیر پارامترهای γ و σ^2 (به ترتیب مقادیر پارامتر تنظیم کننده خطا و پارامتر مربوط به تابع کرنل پایه شعاعی)، در بخش آموزش میزان بایاس مدل (b) و مقادیر α ها که ضرایب لاگرانژ مربوط به تابع می‌باشند تخمین می‌گردد. از نرم افزار GeNIe2.0 جهت شبیه سازی سطح تراز ایستابی با شبکه بیزین استفاده شد. خطای همگرایی ۰/۰۰۱ در نظر گرفته شد.

به منظور مقایسه نتایج حداقل مربعات ماشین بردار پشتیبان و شبکه بیزین در شبیه سازی سطح تراز ایستابی، از شاخص‌های آماری ضریب تبیین (R^2)، ریشه متوسط مربعات خطا (RMSE) و متوسط مطلق خطا (MAE) استفاده شد [۹].

$$R^2 = \left[\frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \cdot \sum_i (y_i - \bar{y})^2}} \right]^2 \quad (16)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (x_i - y_i)^2}{N}} \quad (17)$$

$$MAE = \frac{\sum_{i=1}^N |x_i - y_i|}{N} \quad (18)$$

در روابط ۱۶ الی ۱۸، x_i مقادیر مشاهداتی و y_i مقادیر پیش‌بینی شده و \bar{x} و \bar{y} به ترتیب نشانگر مقادیر میانگین داده‌های متناظر می‌باشند. در پژوهش حاضر ترکیبی که دارای ضریب تبیین بزرگ‌تر، ریشه متوسط مربعات خطا و میانگین خطای مطلق کمتری باشد، به عنوان مدل بهینه انتخاب شد.

نتایج

برای تشخیص ترکیب بهینه متغیرهای ورودی در مدل‌های حداقل مربعات ماشین بردار پشتیبان و شبکه بیزین، ابتدا با استفاده از آزمون گاما، مقدار گاما برای حالت استفاده از تمام متغیرهای ورودی به

در حالت کلی گره‌ها یا والد هستند یا ولد (فرزند) یک گره فرزند می‌تواند توسط چندین والد، تولید شده باشد گره‌هایی که قبل از آن‌ها گره دیگری در گراف وجود دارد با توزیع احتمال شرطی تعریف می‌شوند. در غیر این صورت با احتمال آغازین (اولیه) آنها بیان می‌شوند. احتمالات مربوط به پایین‌ترین بخش در شبکه بیزین از طریق قانون احتمال کل به دست می‌آید و احتمالات مرتبط با بخش‌هایی بالایی شبکه، بر اساس قانون بیزین می‌باشد.

همان‌گونه که در بالا اشاره شد شبکه‌های بیزین امکان محاسبات روبه‌جلو و رو به عقب را برای تحلیل‌گر ایجاد می‌کنند. در واقع نه تنها از تجمیع وضعیت پارامترهای علت می‌توان به وضعیت معلول رسید بلکه در این روش با در اختیار داشتن وضعیت معلول یا همان پارامتر پیش‌بینی شونده با یک فرآیند بازگشت به عقب، امکان محاسبه وضعیت پارامترهای تأثیرگذار به وجود خواهد آمد. به بیان دیگر می‌توان تعیین نمود که میزان تأثیرگذاری هر پارامتر بر ریسک نهایی یا همان خروجی نهایی چه قدر خواهد بود. البته واضح است که اگر مدل شبکه بیزین بر پایه اطلاعات مشاهداتی بسیار اندک توسعه داده و تحلیل شود، خطای تحلیل و پیش‌بینی‌های مربوطه افزایش می‌یابد.

تجزیه و تحلیل داده‌ها

در این تحقیق از داده‌های ماهانه تراز سطح ایستابی چاه‌های پیرومتری دشت اردبیل طی سالهای ۱۳۵۱ الی ۱۳۹۰ استفاده شده است. جهت تعیین کارایی مدل‌های حداقل مربعات ماشین بردار پشتیبان و شبکه بیزین ابتدا تاخیرهای زمانی داده‌ها مشخص گردید. سپس سری داده‌ها به دو قسمت تقسیم‌بندی شدند. ۷۰ درصد برای آموزش مدلها (۱۳۵۱-۱۳۷۸) و ۳۰ درصد برای صحت‌سنجی (۱۳۷۹-۱۳۹۰) استفاده گردید که مشخصات آماری داده‌ها پس از پالایش و حذف داده‌های پرت طبق جدول ۳ می‌باشد. به منظور افزایش کارایی مدل، همه داده‌ها با استفاده از رابطه ۱۵ ابتدا به شکل نرمال بین مقادیر ۱- و ۱ استاندارد شدند [۲۱].

جدول ۴- نتایج آزمون گاما در مدل‌سازی سطح تراز ایستابی دشت اردبیل
 Table 4. Gamma test results in modeling water level plain of Ardabil.

تعداد داده‌ها	نسبت V	خطای استاندارد	شیب (A)	گاما	متغیر غایب در مدل	شماره ترکیب
Data	Ratio	Standard Error	(Gradient(A	Gamma	No variable in the model	Number of combination
480	0081732.0	00040245.0	0.086082	002043.0	-	1
480	011015.0	00045401.0	0070813.0	002753.0	G _(t-6)	2
480	010076.0	00034112.0	0034112.0	002588.0	G _(t-5)	3
480	0099918.0	0036784.0	077074.0	002498.0	G _(t-4)	4
480	0097543.0	0034505.0	082397.0	002498.0	G _(t-3)	5
480	012889.0	00068056.0	059308.0	003224.0	G _(t-2)	6
480	017355.0	00037913.0	043793.0	004378.0	G _(t-1)	7

همان‌طور که در جدول ۵ مشاهده می‌شود، میزان R^2 در بخش آموزش ششمین ترکیب مدل حداقل مربعات ماشین بردار پشتیبان برابر با ۰/۹۷۹ به دست آمده که نسبت به ترکیبات قبلی بزرگ‌تر است. همچنین میزان MAE و RMSE برابر با ۰/۲۰۱ و ۰/۳۰۵ است. در بخش صحت‌سنجی نیز وضعیت به همین ترتیب است. مقدار R^2 در ششمین ترکیب این بخش، بزرگ‌تر از ترکیبات قبلی و برابر با ۰/۹۷۷ است. میزان MAE و RMSE برابر با ۰/۲۰۴ و ۰/۳۰۷ هستند. شاخص‌های آماری یادشده نشان می‌دهند که ترکیب مدنظر نسبت به سایر ترکیبات مدل‌سازی شده با روشهای دیگر و دیگر توابع کرنل از دقت بیشتری برخوردار است. در شکل ۱، مقادیر مشاهداتی و پیش‌بینی‌شده برای ترکیب ششم نشان داده شده است. اصولاً نوع تابع کرنل مناسب باید به روش سعی و خطا تعیین گردد [۱۸]؛ با این حال در اکثر تحقیقات به این موضوع اهمیت داده نشده و تابع کرنل RBF به طور پیش‌فرض به عنوان تابع کاربردی مدل، مورد استفاده قرار گرفته است. با توجه به این که این پیش‌فرض ممکن است در تمام موارد صحیح نباشد، در این تحقیق از روش سعی و خطا استفاده شد. در جدول ۵ ملاحظه می‌گردد که در همه ترکیب‌ها، مدل حداقل مربعات ماشین بردار پشتیبان با تابع کرنل پایه شعاعی نسبت به مدل با توابع کرنل خطی و چندجمله‌ای، مقادیر شبیه‌سازی شده‌ی بهتری ارائه می‌دهد؛ بنابراین می‌توان اطمینان داشت که تخمین ارائه شده توسط تابع کرنل RBF نسبت به توابع دیگر دقیق‌تر است.

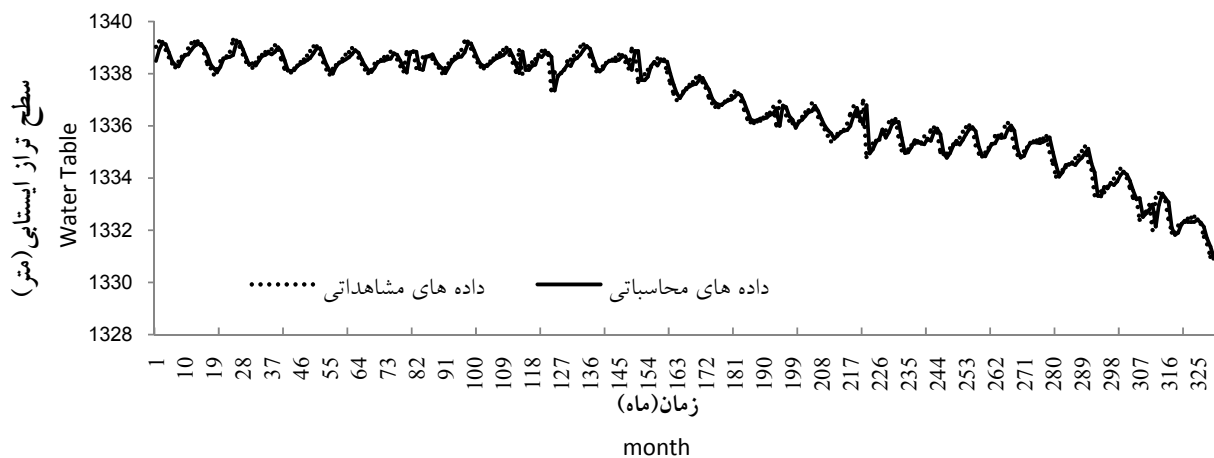
نتایج حاصل از شبیه‌سازی با شبکه بیزین نشان می‌دهد ترکیب ششم در جدول ۵ در بخش آموزش میزان R^2 برابر با ۰/۹۷۹ است که نسبت به ترکیبات قبلی بزرگ‌تر می‌باشد. همچنین میزان MAE و RMSE برابر با ۰/۲۱۵ و ۰/۳۸۸ است. در بخش صحت‌سنجی نیز وضعیت به همین ترتیب است. مقدار R^2 در ششمین ترکیب این بخش، بزرگ‌تر از ترکیبات قبلی و برابر با ۰/۹۷۸ است. میزان

دست آمد. سپس یکی از متغیرهای ورودی حذف و مقدار گاما برای ترکیب مفروض محاسبه شد. آنگاه متغیر مذکور دوباره به مدل وارد و متغیر دیگری حذف شد و مقدار گاما برای ترکیب جدید به دست آمد. این کار برای همه متغیرها تکرار و در جدولی ثبت شد. جدول ۴ نتایج آزمون گاما را برای داده‌های سطح تراز ایستابی دشت اردبیل نشان می‌دهد. در این جدول ترکیب شماره یک، ترکیبی است که در آن همه متغیرهای ورودی حضور دارد. در ترکیبات دیگر همه متغیرها به‌جز یکی در مدل حاضر هستند. مثلاً در ترکیب شماره دو، همه متغیرها به‌جز متغیر $G(t-6)$ حاضر می‌باشد. بعد از اینکه مقدار گاما برای تمامی ترکیب‌های جدول ۴ به دست آمد مقدار آماره (Γ) هر یک از ترکیبات شماره ۲ تا ۷ با مقدار آماره گامای مربوطه به ترکیب ۱ مقایسه شد. ترکیباتی که در آن‌ها مقدار گامای مدل با ترکیبات مفروض بیشتر از مقدار گامای نظیر ترکیب شماره ۱ بود، به‌عنوان مهم‌ترین متغیرها در تخمین سطح تراز ایستابی تشخیص داده شدند و از آن‌ها برای مدل‌سازی استفاده گردید.

به طوری که از جدول ۴ مشاهده می‌شود شش ترکیب وجود دارد که مقدار گامای آن‌ها از ترکیب شماره یک بیشتر می‌باشد. بنابراین متغیرهای غایب این ترکیبات شامل شش متغیر، $G(t-1)$ ، $G(t-2)$ ، $G(t-3)$ ، $G(t-4)$ ، $G(t-5)$ ، $G(t-6)$ به‌عنوان متغیرهای مؤثر در مدل‌سازی سطح تراز ایستابی تشخیص داده شدند. همچنین متغیر $G(t-1)$ به عنوان مهم‌ترین متغیر در این ترکیب می‌باشد. وقتی این متغیر از ترکیب حذف می‌شود مقدار (Γ) نسبت به ترکیب شماره یک (یعنی زمانی که همه متغیرها حضور دارند) بیشترین مقدار را به خود اختصاص می‌دهد. در مرحله بعد بخشی از داده‌ها به منظور مدل‌سازی در دو مدل LS-SVM و BN مورد آموزش قرار گرفت و بخشی نیز جهت صحت‌سنجی مورد آزمایش مدل‌ها قرار گرفت. در جدول ۵ مقادیر شاخص آماری محاسبه‌شده مربوط به نتایج بخش آموزش و صحت‌سنجی در هر دو مدل LS-SVM و BN آورده شده‌اند.

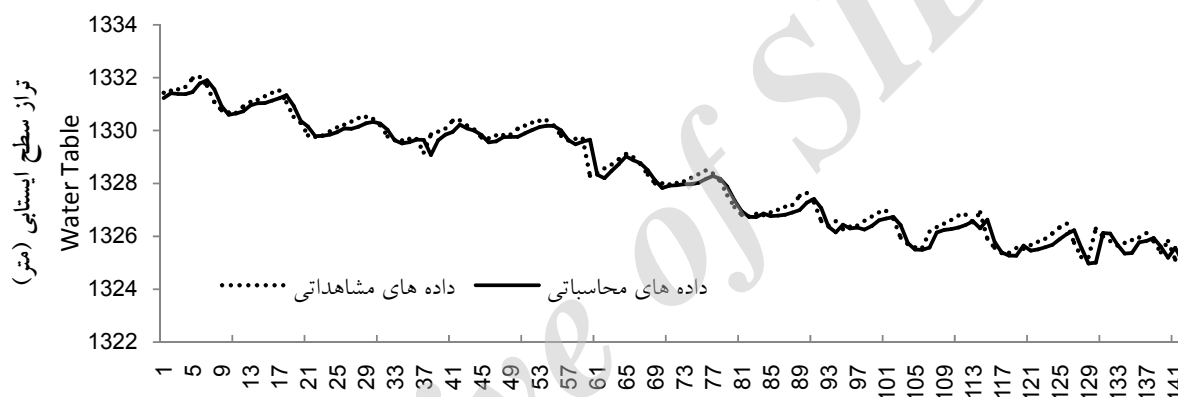
Table 5. Evaluation of Support Vector Machine and Bayesian network models to predict the water level plain of Ardabil

صحت سنجی			آموزش			پارامترهای ورودی Input parameters	مدل Model
Test			Train				
(m)RMSE	MEA	R ²	(m)RMSE	MAE	R ²		
325.0	232.0	958.0	311.0	210.0	975.0	G _{t-1}	LLSVM-RBF
326.0	242.0	954.0	313.0	222.0	973.0	G _{t-1}	LLSVM-Poly
326.0	248.0	957.0	315.0	223.0	972.0	G _{t-1}	LLSVM-Lin
346.0	218.0	976.0	353.0	222.0	975.0	G _{t-1}	Bayesian
313.0	231.0	966.0	310.0	209.0	976.0	G _{t-1} , G _{t-2}	LLSVM-RBF
314.0	234.0	964.0	314.0	215.0	974.0	G _{t-1} , G _{t-2}	LLSVM-Poly
315.0	235.0	963.0	313.0	221.0	973.0	G _{t-1} , G _{t-2}	LLSVM-Lin
340.0	217.0	977.0	349.0	221.0	976.0	G _{t-1} , G _{t-2}	Bayesian
316.0	209.0	963.0	309.0	208.0	978.0	G _{t-1} , G _{t-2} , G _{t-3}	LLSVM-RBF
318.0	215.0	962.0	310.0	219.0	978.0	G _{t-1} , G _{t-2} , G _{t-3}	LLSVM-Poly
319.0	221.0	961.0	312.0	220.0	965.0	G _{t-1} , G _{t-2} , G _{t-3}	LLSVM-Lin
333.0	217.0	977.0	340.0	220.0	976.0	G _{t-1} , G _{t-2} , G _{t-3}	Bayesian
308.0	207.0	966.0	307.0	206.0	977.0	G _{t-1} , G _{t-2} , G _{t-3} , G _{t-4}	LLSVM-RBF
309.0	208.0	966.0	309.0	208.0	975.0	G _{t-1} , G _{t-2} , G _{t-3} , G _{t-4}	LLSVM-Poly
311.0	209.0	965.0	311.0	209.0	967.0	G _{t-1} , G _{t-2} , G _{t-3} , G _{t-4}	LLSVM-Lin
331.0	211.0	978.0	340.0	219.0	977.0	G _{t-1} , G _{t-2} , G _{t-3} , G _{t-4}	Bayesian
308.0	205.0	975.0	307.0	202.0	978.0	G _{t-1} , G _{t-2} , G _{t-3} , G _{t-4} , G _{t-5}	LLSVM-RBF
311.0	213.0	972.0	308.0	211.0	978.0	G _{t-1} , G _{t-2} , G _{t-3} , G _{t-4} , G _{t-5}	LLSVM-Poly
312.0	211.0	974.0	309.0	212.0	967.0	G _{t-1} , G _{t-2} , G _{t-3} , G _{t-4} , G _{t-5}	LLSVM-Lin
329.0	212.0	978.0	337.0	217.0	977.0	G _{t-1} , G _{t-2} , G _{t-3} , G _{t-4} , G _{t-5}	Bayesian
307.0	204.0	977.0	305.0	201.0	979.0	G _{t-1} , G _{t-2} , G _{t-3} , G _{t-4} , G _{t-5} , G _{t-6}	LLSVM-RBF
312.0	206.0	965.0	309.0	204.0	969.0	G _{t-1} , G _{t-2} , G _{t-3} , G _{t-4} , G _{t-5} , G _{t-6}	LLSVM-Poly
315.0	207.0	968.0	307.0	202.0	977.0	G _{t-1} , G _{t-2} , G _{t-3} , G _{t-4} , G _{t-5} , G _{t-6}	LLSVM-Lin
325.0	210.0	978.0	333.0	215.0	979.0	G _{t-1} , G _{t-2} , G _{t-3} , G _{t-4} , G _{t-5} , G _{t-6}	Bayesian



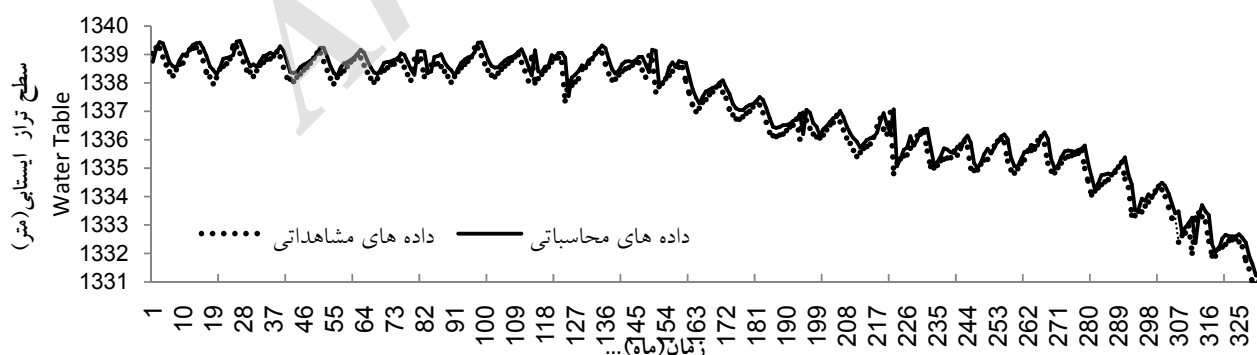
شکل ۱-الف) مقادیر مشاهداتی و پیش‌بینی شده سطح تراز ایستابی دشت اردبیل با بهترین ترکیب حداقل مربعات بردار پشتیبان (ترکیب ۶) بخش آموزش

Fig 1. a. The observed and predicted values of Ardabil water level plain with the best combination of LS-SVM (compound 6) the training



شکل ۱-ب) مقادیر مشاهداتی و پیش‌بینی شده سطح تراز ایستابی دشت اردبیل با بهترین ترکیب حداقل مربعات بردار پشتیبان (ترکیب ۶) بخش صحت سنجی

Fig 1. b. The observed data and predicted water level plain of Ardabil with the best combination of LS-SVM (compound 6) the validation



شکل ۲-الف) مقادیر مشاهداتی و پیش‌بینی شده سطح تراز ایستابی دشت اردبیل با بهترین ترکیب شبکه بیزین (ترکیب ۶) بخش آموزش

Fig 2.a. The observed data and predicted water level plain of Ardabil with the best combination of Bayesian networks (compound 6) the training



شکل ۲-ب) مقادیر مشاهداتی و پیش‌بینی شده سطح تراز ایستابی دشت اردبیل با بهترین ترکیب شبکه بیزین (ترکیب ۶) بخش صحت‌سنجی
 Fig 2.b. The observed data and predicted water level plain of Ardabil with the best combination of Bayesian networks (compound 6) the validation

پیش‌پردازش، می‌توان آزمون گاما را به عنوان روش دقیق و بهینه در جهت کاهش مراحل سعی و خطا در تعیین پارامترهای ورودی در مدل‌سازی ارائه نمود. رضائی و همکاران [۱۷] از روش آزمون گاما جهت پیش‌پردازش داده‌های سطح ایستابی دشت رامهرمز در استان خوزستان جهت مدل‌سازی با ماشین بردار پشتیبان استفاده کردند که نتایج حاصل از عملکرد مناسب آزمون گاما در انتخاب ترکیب بهینه ورودی به مدل می‌باشد. بروجردی و فریدونی [۳] جهت شبیه‌سازی سطح آب زیرزمینی در دشت شیراز از مدل ماشین بردار پشتیبان و شبکه عصبی استفاده کردند. نتایج نشان داد که مدل ماشین بردار پشتیبان از دقت بالایی جهت شبیه‌سازی سطح آب زیرزمینی در منطقه مورد نظر برخوردار است.

منابع

1. Besalsatpour, A. A. Hajabbasi, M. A. and Ayoubi, Sh. 2013. Use of Gamma test technique for choosing the optimum input variable modeling of soil shears strength using artificial neural networks. *Journal of Water and Soil Conservation*. 20(1): 97-114. (In Persian)
2. Biondi, D. and De Luca, D. L. 2012. A Bayesian approach for real-time flood forecasting. *Physics and chemistry of the earth*. 42(44): 91-97.
3. Borojerdi, A. and Fereydoni, M. 2015. Groundwater simulation model using support vector machine and compare it with artificial neural network. *Indian Journal of Fundamental and Applied Life Sciences*. 5(S3): 1368-1375. (In Persian)
4. Durrant, P. J. 2001. Win-Gamma TM: A non-linear data analysis and modeling tool with applications

MAE و RMSE برابر با $0.3010/210$ هستند. شاخص‌های آماری یادشده نشان می‌دهند که ترکیب مدنظر نسبت به سایر ترکیبات مدل‌سازی شده با این روش از دقت بیشتری برخوردار است. در شکل ۲، مقادیر مشاهداتی و پیش‌بینی شده برای ترکیب ششم نشان داده شده است.

بحث و نتیجه‌گیری

سطح آب زیرزمینی یکی از مهم‌ترین فاکتورها در طرح‌های عمرانی و کشاورزی بوده و شبیه‌سازی برآورد دقیق آن از اهمیت زیادی برخوردار است. در این مطالعه، از روش حداقل مربعات ماشین بردار پشتیبان و شبکه بیزین برای تخمین سطح تراز ایستابی دشت اردبیل استفاده شد. نتایج حاصل از پژوهش نشان داد که برای شبیه‌سازی سطح تراز ایستابی، ساختاری که در آن ورودی مدل از ترکیب پارامترهای سطح تراز ایستابی با شش تأخیر زمانی تشکیل شده باشد از دقت بالاتری برخوردار است. همچنین با توجه به معیارهای ارزیابی برای سنجش مدلها، نتایج نشان داد که هر دو روش مورد بررسی، می‌تواند با دقت نسبتاً بالایی سطح تراز ایستابی را پیش‌بینی نماید. در این میان، مدل حداقل مربعات ماشین بردار پشتیبان با توجه به خطای کمتر نسبت به مدل شبکه بیزین عملکرد بهتری از خود نشان داده است. نتایج مدل‌سازی ترکیب ششم، بیان‌گر کارایی پیش‌پردازش آزمون گاما در کاهش مراحل سعی و خطا و ارائه ترکیب بهینه بوده است. مهم‌ترین مزیت مدل شبکه بیزین آن است که با توجه به رابطه‌ای که داده می‌شود می‌تواند برای پیش‌بینی سطح تراز ایستابی در سالهای آتی استفاده گردد. با توجه به نتایج به‌دست‌آمده در بخش مدل‌سازی، می‌توان مدل حداقل مربعات ماشین بردار پشتیبان را به عنوان مدلی با دقت و سرعت پردازش بالا در مدل‌سازی مسائل هیدرولوژیکی و مسائلی با پارامترهای ورودی زیاد تأثیرگذار بر پارامتر خروجی، پیشنهاد نمود. همچنین با توجه به نتایج به‌دست‌آمده از بخش

14. Moghaddamnia, A. Ghafari Gousheh, M. Piri, J. Amin, S. and Han, D. 2009. Evaporation estimation using artificial neural networks and adaptive neuro fuzzy inference system techniques. *Advance in Water Resource*. 32: 88-97. (In Persian)
15. Madadgar, Sh. and Moradkhani, H. 2014. Spatio-temporal drought forecasting within Bayesian networks. *Journal of Hydrology*. 512: 134-146. (In Persian)
16. Remesan, R. Shamim, M.A. and Han, D. 2008. Model data selection using gamma test for daily solar radiation estimation. *Hydrological Processes*. 22: 4301-4309.
17. Rezaee, E. Shahidi, A. Khashee, A. and Riyahi Madvar, H. 2014. Application of Least Squares Support Vector Machine Model for Water Table Simulation (Case Study: Ramhormoz plain). *Iranian Journal of Irrigation and Drainage*. 4:510-520. (In Persian)
18. Seifi, A. 2010. Developing an Expert System for Predicting Daily Reference Evapotranspiration Using Support Vector Machines (SVM) in Comparison with ANFIS, ANN and Empirical Methods. M. S. in Irrigation and drainage Engineering, Tarbiat Modares University, Tehran. (In Persian)
19. Seifi, A. and Riahi-Madvar, H. 2012. Input Variable Selection in expert systems based on hybrid Gamma Test-Least Square Support Vector Machine, ANFIS and ANN models. Provisional Chapter. INTECH.
20. Sharifi, A. Dinpashoh, Y. Fakhri-Fard, A. and Moghaddamniya, A. 2013. Optimum Combination of Variables for Runoff Simulation in Amameh Watershed using Gamma test. *Journal of Water and Soil* 23(4): 59-72. (In Persian)
21. Xu, L., J. Wang, J. Guan, and F. Huang. 2007. A Support Vector Machine Model for Mapping of Lake Water Quality from Remote-Sensed Images. *IC-MED*, 1(1): 57-66.
- to flood prediction. PhD Thesis, Department of Computer Science, Cardiff University, Wales, UK.
5. Davies, P. 2007. Bayesian decision networks for management of high conservation. Report to the Conservation of Freshwater Ecosystem Values Project. Department of Primary Industries and Water. Hobart Tasmania.
6. Evans, D. and Jones, A.J. 2002. A proof of the gamma test. Pp. 2759-2799. *Proceedings of Royal Society. Series A* 458 (2027).
7. Farmani, R. Henriksen. H.J. and. Savic. D. 2009. An evolutionary Bayesian belief network methodology for participatory decision making under uncertainty: an application to groundwater management *Integrated Environmental Assessment and Management*, 8(3):456-61.
8. Ghabaei Sough, M. Mosaedi, A. and Dehghani, A.A. 2011. Solar radiation data and their intelligent modeling based on gamma test with evaluation of calibrated empirical equations. *Journal of Water and Soil Conservation*, 18(4):185-208. (In Persian)
9. Ghorbani, M.A. and Dehgani, R. 2015. Application of Bayesian Neural Networks, Support Vector Machines and Gene Expression Programming Analysis of Rainfall - Runoff Monthly (Case Study: Kakarza River). *Journal of Irrigation Engineering University of Ahvaz*. 39(20):125-138. (In Persian)
10. Izadi, A. Davari, K. Alizadeh, A and B. Ghahreman. 2008. Application of Panel Data Model in Predicting Groundwater Level. *Iranian Journal of Irrigation & Drainage*. 2: 59-71. (In Persian)
11. Jones AJ, 2004. New tools in non-linear modeling and prediction. *Computational Management Science*. 1: 109-149.
12. Mellit, A. Massi Pavan, A. and Benganem, M. 2013. Least squares support vector machine for short-term prediction of meteorological time series. *Theoretical and Applied Climatology*. 111: 297-307.
13. Mirzaei, A.A. and Nazemi, A.H. 2011. Prediction of water table level using intelligent systems. *Journal of Water Engineering*, 4:1-10. (In Persian)

Abstract

Application of Bayesian Networks and Support Vector Machine Model to Predict Changes in Water Level (Case Study: Ardebilplain)

F. Akhoni Pourousseini¹ and E. Asadi²

Received: 2016/02/09 Accepted: 2016/11/21

Groundwater has been raised as one of the major sources of water supply for drinking and agriculture, especially in arid and semi-arid. Simulation of groundwater system because of the complexity of these systems is a difficult task. In this paper, using data Ardebil plain water level in the period (1972-2011), the evaluation and selection of appropriate inputs for processing gamma test performance and efficiency of the least squares support vector machines and Bayesian network models were discussed. Monthly water level as input parameters with different delays gamma test was considered. Gamma test results showed that the water level by 6 latency, offers better results to predict. Water level simulation using least squares support vector machines and Bayesian network models also showed that the input structure to predict the water level the next month will be delayed until six. The two models with the same input structure, least squares support vector machine model, better performance, according to the coefficient of determination 0.977, mean absolute error 0.204 and root mean square error 0.307, compared to Bayesian networks have. The results showed that gamma test compound in the appropriate input soft computing can have a better performance.

Keywords: Bayesian networks, Gamma Test, Groundwater level, Plain Ardebil, Support Vector Machine

1- Ms.c student of Water Resources Management, University of Tabriz, Corresponding Author, Email: fateme.pourhosseini@yahoo.com

2- Assistant Professor of Water Engineering, University of Tabriz, Tabriz, Iran, Email: esasadi@gmail.com