مرتع و آبخیزداری، مجله منابع طبیعی ایران
دوره ۷۳، شماره ۱، بهار ۱۳۹۹
ص ۱۸-۱

تاریخ دریافت: ۱۳۹۶/۱۱/۰۲
تاریخ تصویب: ۱۳۹۷/۱۱/۱۸

# پیش‌بینی سلامت محیطی با استفاده از برنامه‌ریزی بیان ژن و شبکهٔ بیزین در حوزهٔ آبخیز طالقان

❖ **پیام ابراهیمی؛** دانشجوی دکتری دانشکدهٔ منابع طبیعی، دانشگاه تهران، کرج، ایران.

❖ **علی سلاجقه*؛** استاد دانشکدهٔ منابع طبیعی، دانشگاه تهران، کرج، ایران.

❖ **محسن محسنی ساروی؛** استاد دانشکدهٔ منابع طبیعی، دانشگاه تهران، کرج، ایران.

❖ **آرش ملکیان؛** دانشیار دانشکدهٔ منابع طبیعی، دانشگاه تهران، کرج، ایران.

❖ **امیر سعدالدین؛** دانشیار دانشکدهٔ مرتع و آبخیزداری، دانشگاه علوم کشاورزی و منابع طبیعی گرگان، گرگان، ایران.

**چکیده**

یکی از معیارهای مهم کیفیت زندگی، سلامت حوزهٔ آبخیز است. پژوهش‌های صورت گرفته در این زمینه نشان می‌دهد در ایران مدلی برای بررسی سلامت حوزهٔ آبخیز تهیه و تنظیم نشده است. در این پژوهش با استفاده از طول آماری ۲۷ ساله (۱۳۶۹-۱۳۹۵) ۵ متغیر محیطی (رسوب، دبی، بارش، دما، تبخیر) در حوزهٔ آبخیز طالقان واقع در استان البرز با استفاده از بیان ژن و شبکهٔ بیزین برآورد شده است. با استفاده از برنامه‌ریزی بیان ژن و شبکهٔ بیزین هر متغیر، دورهٔ ۱۳۸۴- ۱۳۶۹ به‌عنوان دورهٔ آموزش و دورهٔ ۱۳۹۲-۱۳۸۵ به‌عنوان دورهٔ آزمون و دورهٔ ۱۳۹۵-۱۳۹۳ به عنوان دورهٔ اعتبار سنجی (میزان دقت پیش‌بینی) مدل انتخاب شد. در مقایسه دقت برآوردی مدل برنامه‌ریزی بیان ژن و شبکهٔ بیزین به ترتیب مقادیر همبستگی میانگین پنج متغیر ۰/۸۷ و ۰/۷۸ می‌باشد. در مورد مدل بیان ژن مقادیر ضریب تعیین در بخش آموزش به ترتیب دبی ۰/۸۷، رسوب ۰/۹۲، بارش ۰/۸۹، دما ۰/۹۱ و تبخیر ۰/۷۷ و همچنین در مدل بیزین مقادیر به ترتیب ۰/۷۳، ۰/۸۸، ۰/۷۸، ۰/۸۱ ۰/۷۴ می‌باشد. با توجه به الگوی فضایی پراکنش داده‌ها مقادیر برنامه‌ریزی بیان ژن با توجه به تولید نسل ۲۰۰ هزار بار قدرت بالایی در شبیه‌سازی مقادیر آتی خواهد داشت. نتایج این پژوهش نشان می‌دهد وضعیت سلامت حوزهٔ آبخیز با کسب امتیاز ۸ در سال ۱۳۹۵ وضعیت سرطان پیشرفته را دارد و با توجه به نتایج مدل در سال ۱۳۹۶ نیز در بهترین حالت می‌تواند در وضعیت سرطان عود کننده قرار گیرد.

**کلید واژگان:** سلامت حوزهٔ آبخیز، طالقان، بیان ژن، شبکهٔ بیزین.

* نویسنده مسئول: شماره تماس: ۹۸۹۱۲۳۸۴۲۴۶۴+
Email: Salajegh@ut.ac.ir

## ۱. مقدمه

امروزه این حقیقت به خوبی دریافت شده است که سیستم‌های انسانی و محیط جغرافیایی در ارتباط تنگاتنگی با یکدیگر هستند و سلامت یک سیستم بر پایداری سیستم دیگر و پایداری یک سیستم بر سلامت سیستم دیگر تأثیر به سزایی دارد [۲]. بررسی سلامت و پایداری محیط‌زیست و به ویژه سیستم‌های اکولوژیک بر اقتصاد اکولوژیک انسان بسیار اثرگذارند چرا که منابع تجدیدپذیر و تجدیدناپذیر که جزء سرمایه‌های طبیعی محسوب می شوند با اثرات خلاقانه سیستم‌های انسانی دائماً در حال تغییر هستند و طبق قوانین فیزیک از شکلی به شکل دیگر از انرژی تبدیل می‌شوند [۱۷]. تلاش‌های بسیاری در علوم فرابخشی به ویژه ادغام سیستم‌های انسانی و محیطی و تأثیر این دو بر یکدیگر صورت گرفته تا با توجه به مقیاس‌های زمانی، مکانی و سازمانی متفاوت این دو سیستم ساخت‌های مفهومی و بنیادینی شکل گیرد تا چالش‌های دنیای واقعی بررسی شود اما نگاه تک‌جانبه گرایانه تعادل بین انسان و محیط، بخشی را ضعیف‌تر از بخش دیگر جلوه می‌دهد [۵۱، ۵۹]. علاوه بر تنازع بقا در میان موجودات اکوسیستم محیطی، تنازعاتی در بهره‌گیری از دو عنصر ویژهٔ آب و زمین در کل جهان هستی وجود دارد. محدودیت منابع طبیعی در تقابل نیازهای نامحدود بشری، لزوم مدیریت جامع منابع طبیعی و انسان را بیش از پیش ملموس‌تر می‌سازد [۴ و ۷]. در حال حاضر در بسیاری از جوامع انسانی سیستم مدیریت محیط زیست (EMS[1])، به‌عنوان برترین سیستم مدیریتی به منظور مراقبت از محیط زیست بر پایهٔ بررسی آثار محیطی مورد استفاده قرار می‌گیرد [۳۷ و ۴۷]، اما تأثیراتی که فعالیت‌های انسانی بر کیفیت آب، میزان سیل، افزایش دمای کرهٔ زمین و یا در مجموع بر پیکرهٔ منابع طبیعی می‌گذارد نیازمند کمی‌سازی و پیش‌بینی است که در این زمینه بیان ژن یکی از محبوب‌ترین روش‌های پیش‌بینی در سال‌های اخیر است [۵۰]. به‌منظور بررسی سلامت محیط در ارتباط با انسان نیاز به دو اصل اساسی است، ابتدا تعریف موقعیت جغرافیایی مشترک بین انسان و محیط و دیگری یک تئوری آماری و ریاضی منطقی از طریق برنامه‌ریزی ژنتیک و الگوریتم ژنتیک بین این دو محیط است [۴۱]. انواع روش‌های آماری و ریاضی برای تعیین آستانهٔ سلامت و پایداری محیط توسعه یافته‌اند اما این مدل‌ها یا منحصر به منطقه‌ای خاص هستند و یا یک وضعیت خاص را در نظر می‌گیرند که با قوانین پویای طبیعت و انسان در تعارض است و جنبهٔ تئوریک آن با فرض‌های بسیاری روبه رو است که در طبیعت و رفتار انسان صادق نیست [۲۸، ۸، ۲۲، ۵۲، ۶۰ و ۱۲ ] و یک راهنمای صحیحی از مدل‌های بررسی سلامت محیط در دسترس نیست [۳۸، ۵۵، ۱۱]. اما می‌توان از بیان ژن در پیش‌بینی داده های محیطی استفاده نمود. محققانی کاربرد مدل‌های ترکیبی هوش مصنوعی-موجک را مورد بررسی قرار دادند. نتایج این بررسی نشان داد که علت مورد توجه قرار دادن تبدیلات موجک، به خاطر فایده و سودمندی تجزیه و تحلیل تفکیک‌پذیری چند جانبه، حذف اختلالات مربوط به امواج الکتریکی یا الکترومغناطیسی و نیز قابلیت قدرتمند هوش مصنوعی در بهینه‌سازی، تطبیق پذیری و تخمین فرآیندهاست [۴۶]. در تحقیقی یک مدل ترکیبی بیان ژن-موجک برای پیش‌بینی کوتاه مدت و بلند مدت دمای هوا ارائه شد [۲۳]. این مطالعه با استفاده از داده‌های روزانه و ماهانهٔ دما در ایستگاه‌های مهاباد و ارومیه صورت گرفت. نتایج بیانگر قابلیت مدل بیان ژن بوده است. در پژوهشی از مدل ترکیبی بیان ژن-موجک در پیش‌بینی رواناب با استفاده از داده‌های بارش رواناب استفاده شد [۵۳]. در این مطالعه از توابع موجک مختلف استفاده و بهترین تابع انتخاب گردید. محققینی از مدل ترکیبی بیان ژن و موجک برای پیش‌بینی کوتاه مدت و بلند مدت جریان رودخانه‌ای در ترکیه استفاده کردند [۱۹]. نتایج

[1] Environment Management System

بیانگر عملکرد بهتر مدل ساده بود. عدم قطعیت در طبیعت، کمبود و یا نواقص آماری، تقابل سیاست‌های اقتصادی و اجتماعی انسانی، عدم هماهنگی میان مقیاس مکانی و زمانی اکوسیستم‌های انسانی و محیطی از جمله عواملی است که باعث می‌گردد مدیریت جامعی میان این دو اکوسیستم برقرار نگردد و مدیریت پایدار و کارآمد نیز چارچوب مشخصی را سازماندهی ننماید [۲۴]، اما این ناکارآمدی‌ها نباید مانع از مدیریت این مجموعه گردد و باید اثرات کنشی انسان بر حوزهٔ آبخیز به عنوان عرصهٔ فعالیت و محیط زیست کمی شود. در برخی شاخص‌های حیاتی مانند کیفیت آب، حجم رواناب، میزان رسوب می‌توان اثرات انسان و وضعیت حوزهٔ آبخیز را بررسی و پیش‌بینی نمود و بر اساس آن فعالیت‌های حمایت تصمیم چند معیاره (MCDS[1]) در برنامه‌ریزی استراتژیک محیط زیستی را طراحی نمود [۲۷، ۱۸، ۱، ۲۱، ۴۲، ۴۴ و ۴۵]. انتخاب متغیرهای مؤثر بر سلامت محیط همواره یکی از مهم‌ترین بخش‌های ارزیابی وضعیت سلامت محیطی است. در این پژوهش با استفاده از اطلاعات قابل استناد دستگاه‌های اجرائی مهم‌ترین متغیرهای مؤثر بر سلامت تعین شد و سپس بر مبنای داده‌کاوی اطلاعات موجود وضعیت احتمالی سال ۱۳۹۶ حوزهٔ آبخیز طالقان مشخص شده است تا بر مبنای آن اطلاعات دقیقی از روند تغییرات مشخص شود.

## ۲. روش‌شناسی

### ۱،۲. معرفی منطقهٔ مورد مطالعه

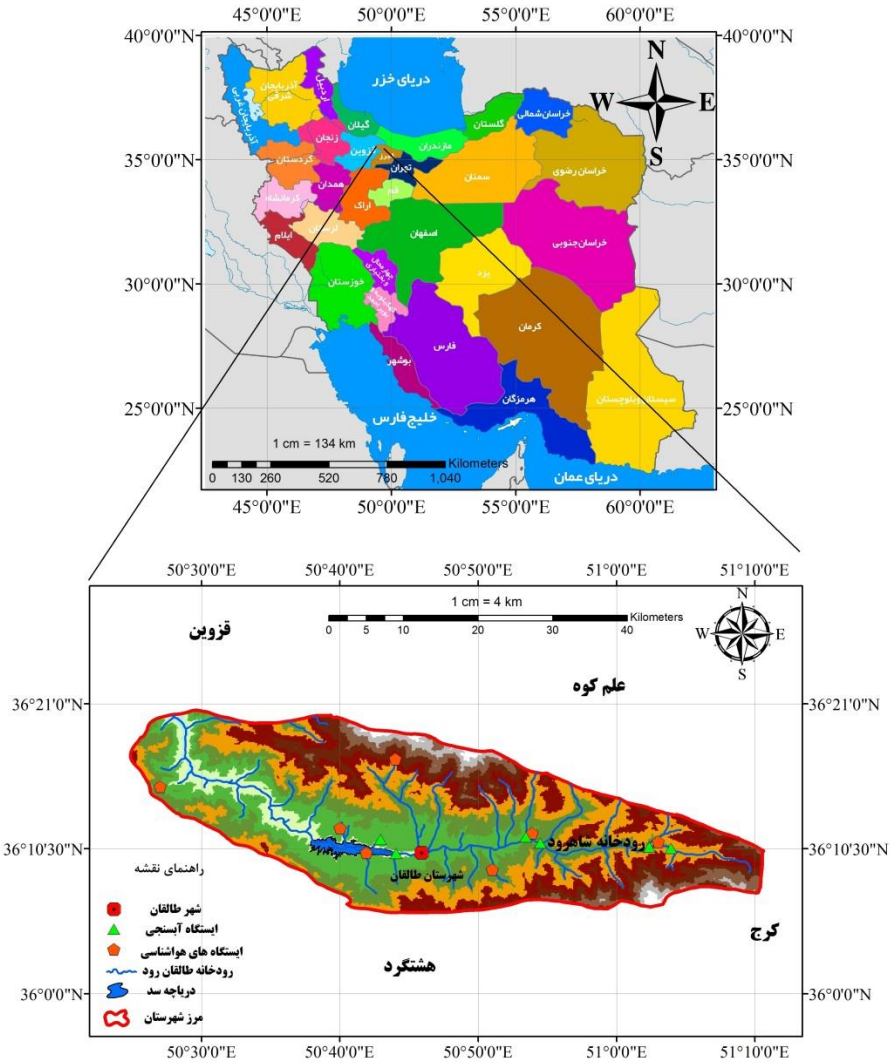
حوزهٔ آبخیز طالقان یکی از شاخه‌های آبخیز سفیدرود به شمار می‌رود که در دامنهٔ جنوبی رشته‌کوه‌های البرز در شمال غرب کرج واقع شده است. این حوزهٔ آبخیز در بین عرض‌های شمالی ۳۶ درجه و ۲۰ دقیقه و ۴۸ ثانیه و ۳۶ درجه و ۵ دقیقه و ۲۳ ثانیه و طول‌های شرقی ۵۰ درجه و ۳۹ دقیقه و ۳۵ ثانیه و ۵۱ درجه و ۱۱ دقیقه و ۷ ثانیه واقع شده است. ین منطقه با وسعتی حدود ۱۱۱ هزار هکتار در فاصلهٔ ۱۰۵کیلومتری مرکز استان (کرج) و ۱۲۰ کیلومتری شمال غربی تهران واقع شده است و عرصه‌های ملی شهرستان ۸۹ هزار هکتار است که از این مساحت ۴۸۰۷ هکتار جنگل و ۸۴۶ هزار هکتار مرتع می‌باشد. این شهرستان شامل ۲ بخش، ۴ دهستان و ۷۵ روستا می باشد. نوع اقلیم براساس روش دومارتن مدیترانه‌ای تا مرطوب و ارتفاع از سطح دریا از ۱۲۶۰ تا ۴۱۲۵ متر متغیر است. رودخانهٔ طالقان یا طالقان رود، رودخانهٔ مهم شهر طالقان است که از کوه‌های کندوان و کهار بزرگ سرچشمه می‌گیرد. ۶۳ درصد از مساحت این شهرستان نیز به واسطهٔ وجود ارتفاعات و شیب‌های منتهی به سد طالقان به عنوان یکی از منابع مهم تأمین آب در زمرهٔ اراضی حفاظتی قرار گرفته‌اند. از طرفی زمین‌های دارای توان برای توسعهٔ سکونت و صنعت در این شهرستان بسیار ناچیز است و نمی‌توان برای این شهرستان توسعهٔ سکونت و صنعت را بیش‌ازحد شرایط موجود تصور کرد. در شکل (۱)، موقعیت منطقهٔ مورد مطالعه نشان داده شده است. همچنین در جدول (۱) نام و موقعیت ایستگاه‌های مورد بررسی در حوزهٔ آبخیز طالقان آورده شده است.

### ۲،۲. روش تحقیق

داده های مورد استفاده در این تحقیق، برای دورهٔ آماری (۱۳۹۵-۱۳۶۹) به سه مجموعه، داده‌های آموزش مدل (۱۳۸۴-۱۳۶۹) و داده های آزمون مدل (۱۳۹۲-۱۳۸۵) و داده های پیش‌بینی مدل (۱۳۹۵-۱۳۹۳). با توجه به پژوهش‌هایی که در این زمینه انجام شده است نسبت داده‌های آموزش ۶۰ تا ۷۰ درصد سال‌های مورد بررسی و نسبت داده‌های آزمون ۳۰ تا ۴۰ درصد داده‌های آموزش می‌باشد [۳] . در مورد سال‌های مورد پیش‌بینی با توجه به مدل‌سازی سری داده به صورت x3=x2+x1 خواهد بود که در آن x1 داده‌های سال گذشته، x2

[1] Multi-Criteria Decision Support

داده‌های سال جاری و x3 مقادیر مورد پیش‌بینی در سال آتی می‌باشد. لذا مدل نیاز به دو سال آمار جهت ورودی خواهد داشت. برنامه‌ریزی بیان ژن که در ادامهٔ سیر تکاملی مدل‌های هوشمند به وجود آمده و به ابزارهای مدل‌سازی مسائل مربوط به تعیین ساختار پدیده‌ها افزوده شده، جزء روش‌های الگوریتم گرد شی محسوب می شود که مبنای تمامی آن‌ها بر اساس نظریهٔ تکامل داروین استوار است. مزیت برنامه‌ریزی بیان ژن نسبت به مدل‌های دیگر از جمله شبکه‌های عصبی این است که در برنامه‌ریزی بیان ژن، ابتدا ساختار بلوک‌ها (متغیرهای وردی، هدف و مجموع توابع) تعریف شده و سپس ساختار بهینهٔ مدل و ضرایب طی فرآیند آموزش تعیین می شوند، در حالی که در شبکه‌های عصبی، ابتدا ساختار تعیین شده و فقط ضرایب مدل، طی فرآیند آموزش حاصل می‌شوند. همچنین این الگوریتم به طور خودکار می‌تواند متغیرهای ورودی که در مدل بیشترین تأثیر را دارند انتخاب کند. این بخش با استفاده از ۵ شاخص محیطی سلامت حوزهٔ آبخیز طالقان را مورد سنجش قرار داده است. این مدل برای اولین بار در ایران به صورت بومی توسط دانشکدهٔ منابع طبیعی دانشگاه تهران در سال ۱۳۹۶ توسط نگارندگان این پژوهش تهیه و تنظیم شده است.




**شکل ۱. موقعیت منطقهٔ مورد مطالعه**

**جدول ۱. ایستگاه‌های مورد استفاده در حوزهٔ آبخیز طالقان**

| ردیف | نام ایستگاه | نام رودخانه | نوع ایستگاه | سال تأسیس | مختصات جغرافیایی | | ارتفاع |
|---|---|---|---|---|---|---|---|
| | | | | | عرض | طول | |
| ۱ | گته ده | گته ده | آبسنجی | ۱۳۷۵ | ۵۱□۴'۱۱" | ۳۶□۱۰'۱۹" | ۲۳۲۲ |
| ۲ | دهدر | دهدر | آبسنجی | ۱۳۷۵ | ۵۱□۲'۱۹" | ۳۶□۱۰'۴۰" | ۲۲۸۶ |
| ۳ | جوستان | آلیزان | آبسنجی | ۱۳۶۹ | ۵۰□۵۳'۲۹" | ۳۶□۱۱'۱۷" | ۱۹۵۶ |
| ۴ | جوستان | شاهرود | آبسنجی | ۱۳۷۰ | ۵۰□۵۴'۲۴" | ۳۶□۱۰'۵۵" | ۱۹۵۲ |
| ۵ | جوستان | مهران | آبسنجی | ۱۳۷۱ | ۵۰□۵۴'۳۷" | ۳۶□۱۰'۵۳" | ۱۹۸۸ |
| ۶ | گلینک | شاهرود | آبسنجی | ۱۳۳۷ | ۵۰.۴۳۳۸۸۹ | ۳۶.۱۶۸۶۱۱ | ۱۷۷۰ |
| ۷ | انگه | * | بارانسنج معمولی | ۱۳۷۴ | ۵۰□۲۷'۴۶" | ۳۶□۱۵'۳۰" | ۱۹۱۷ |
| ۸ | گته ده | * | بارانسنج معمولی | ۱۳۷۴ | ۵۱□۴'۱۲" | ۳۶□۱۰'۲۳" | ۲۳۰۴ |
| ۹ | دهدر | * | بارانسنج معمولی-ثبات | ۱۳۷۴ | ۵۱□۲'۵۱" | ۳۶□۱۱'۲۱" | ۲۳۲۵ |
| ۱۰ | سگرانچال | * | بارانسنج معمولی | ۱۳۷۴ | ۵۰□۴۹'۲۹" | ۳۶□۱۲'۲" | ۲۲۰۲ |
| ۱۱ | آرموت | * | بارانسنج معمولی-الکترونیکی | ۱۳۷۴ | ۵۰□۴۰'۱۳" | ۳۶□۱۱'۲۳" | ۱۸۰۳ |
| ۱۲ | جوستان | * | بارانسنج الکترونیکی | - | ۵۰□۵۳'۲۸" | ۳۶□۱۱'۱۹" | ۱۹۴۸ |
| ۱۳ | جوستان | * | تبخیرسنجی درجه ۱ | ۱۳۷۱ | ۵۰□۵۳'۴۴" | ۳۶□۱۱'۲۲" | ۱۹۸۳ |
| ۱۴ | زیدشت | * | تبخیرسنجی درجه ۱ | ۱۳۴۵ | ۵۰□۴۱'۵۴" | ۳۶□۱۰'۱۳" | ۱۷۵۱ |

بر اساس تحقیقات مختلف [۳۰، ۵، ۶۳، ۱۴، ۹، ۴۹، ۶۴] متغیرهای سلامت محیطی به ترتیب رسوب تولیدی، دبی، دما، تبخیر و بارش می‌با شد که در جدول (۲) آورده شده است. متغیر رسوب با توجه به سازند زمین شناسی حوزهٔ طالقان از اولویت بالاتری برخوردار است زیرا تحقیقات در این حوزهٔ آبخیز عامل رسوب تولیدی و پتانسیل فرسایش پذیری را به‌گونه‌ای با اهمیت می‌داند که در وضعیت کنونی بالاترین اولویت حفاظت خاک خواهد بود [۵۶، ۶]، از طرفی موقعیت سد طالقان که تمامی آب‌های خروجی از ایستگاه گلینک را در معرض تبخیر قرار می‌دهد در چرخهٔ هیدرولوژیک مؤثر است و تحقیقات نشان می‌دهد وجود سد در منطقه بازدهی آب جمع‌آوری شدهٔ مخزن را به شدت کاهش می‌دهد و با توجه به این‌که ایران به لحاظ کمبود مقادیر بارش این متغیر نقش حیاتی دارد [۲۰، ۱۶، ۳۶،۳۵،۳۴،۳۳،۳۲،۳۱، ۴۰] تبخیر مخازن آب نظیر سد، دریاچه که حجم بالایی از آب را در خود نگه می‌دارند ممکن است نتایجی نظیر دریاچهٔ ارومیه را به ارمغان آورد. اگرچه عوامل دیگری در تبخیر دریاچهٔ ارومیه مؤثر بوده است اما فرآیند تبخیر از سطح دریاچه است که مکانیسم را ایجاد کرده و ادامه می‌دهد. براساس پژوهش‌های صورت گرفته [۲۵]، مدل بیان ژن ساده با مقادیر همبستگی بالاتر نسبت به سایر مدل‌های ترکیبی از دقت بالاتری برخوردار بوده است. مدل‌های مذکور براساس نتایج مدل بیان ژن بوده و داده‌کاوی صورت گرفته قابل تعمیم به تمام داده‌هایی است که هدف از بررسی داده‌ها کاوش و یافتن فرآیند درونی آن‌هاست و محدود به رشته‌ای خاص نخواهد بود [۲۶، ۴۳، ۵۷]. قطعاً مدل‌هایی نظیر بیزین، شبکهٔ عصبی، تئوری پروانه، بهینه سازی هوشمند، محاسبات تکاملی از درون علم محیط زیست یا منابع طبیعی برنخواسته‌اند که قابل تعمیم نباشند، لذا بدون ترکیب مدل با سایر مدل‌های هیبریدی به مقایسهٔ نتایج به دست آمده این مدل با مدل شبکهٔ بیزین پرداخته شد تا علاوه بر بررسی نتایج مدل، وضعیت سلامت سال‌های ۱۳۹۵ و ۱۳۹۶حوزهٔ

آبخیز طالقان را بر اساس متغیرهای محیطی برآورد نماییم.

### ۳،۲. شبکهٔ بیزین[1]

این روش یکی از روش‌های سیستم پشتیبان تصمیم‌گیری می‌باشد که ابزار قدرتمندی در مدل کردن روابط علی و معلولی در قالب شبکه‌ای از احتمالات است. نکتهٔ بسیار مهم در مورد شبکهٔ عصبی بیزین این است که این روش به اطلاعات دقیق و تاریخچهٔ کامل یک واقعیت نیاز ندارد بلکه می‌تواند با استفاده از اطلاعات ناقص و غیر دقیق نیز به نتایج بسیار قانع کننده‌ای در زمینهٔ تخمین وضعیت فعلی یا آیندهٔ یک سیستم دست یابد. همچنین یک روش سازگار و انعطاف پذیر برای مدل کردن موقعیت‌های نامطمئن و غیر قطعی براساس روابط علت و معلولی است. در شبکه‌های عصبی یکی از روش‌های دست‌یابی مؤثر به پارامترهای بهینه، استفاده از تنظیم بیزین است که به طور اتوماتیک مقادیری مناسب برای پارامتر تابع قرار می‌دهد. روش تنظیم بیزین می‌توانند جهت بهبود توانایی شبکهٔ عصبی و آموزش تابع هدف F که به صورت زیر نشان داده شده است استفاده شود.

$$F = \beta E_D + \alpha E_W$$ فرمول (۱)

که در آن EW مجموع مربعات وزن‌های شبکه و ED مجموع مربعات باقی‌مانده بین پاسخ شبکه و تابع هدف می‌باشد. α و β نیز پارامترهای تابع هدف می‌باشند. که هر یک از این پارامترها به آموزش شبکه در کاهش باقی ماندهٔ خروجی‌ها یا حجم شبکه بستگی دارد. نکتهٔ اساسی روش تنظیم این است که پارامترهای تابع هدف را چگونه از طریق داده‌های آماری بیزین انتخاب و بهینه نماید. به‌طوری که اگر α و β به عنوان متغیرهای تصادفی در نظر گرفته شود، این روش یکی از روش‌های سیستم پشتیبان تصمیم‌گیری می‌باشد که ابزار قدرتمندی در مدل کردن روابط علی و معلولی در قالب شبکه‌ای از احتمالات است. نکتهٔ بسیار مهم در مورد شبکهٔ عصبی بیزین این است که این روش به اطلاعات دقیق و تاریخچهٔ کامل یک واقعیت نیاز ندارد بلکه می‌تواند با استفاده از اطلاعات ناقص و غیر دقیق نیز به نتایج بسیار قانع کننده‌ای در زمینهٔ تخمین وضعیت فعلی یا آیندهٔ یک سیستم دست یابد. همچنین یک روش سازگار و انعطاف پذیر برای مدل کردمن موقعیت‌های نامطمئن و غیر قطعی براساس روابط علت و معلولی است. در شبکه‌های عصبی یکی از روش‌های دست‌یابی مؤثر به پارامترهای بهینه، استفاده از تنظیم بیزین است که به طور اتوماتیک مقادیری مناسب برای پارامتر تابع قرار می‌دهد. روش تنظیم بیزین می‌تواند جهت بهبود توانایی شبکهٔ عصبی و آموزش تابع هدف F که به صورت زیر نشان داده شده است، استفاده شود.

قانون بیزین به صورت زیر فرموله می‌شود:

$$P(\propto, \beta|D, M) = \frac{P(D|\alpha, \beta, M)P(\alpha, \beta|M)}{P(D|M)}$$ فرمول (۲)

در فرمول فوق D داده‌های آموزش، M مدل شبکه و W نیز وزن شبکه می‌باشد. بر طبق قانون بیزین اگر α و β از توزیع یکنواخت پیروی کنند در این صورت احتمال $P(D|\propto, \beta, M)$ زمانی ماکزیمم می‌شود که احتمال توزیع اولیه α و β در معادلهٔ فوق حداکثر به مقدار ماکزیمم خواهد رسید. با فرض این که باقی مانده‌ها و وزن‌ها متغیرهای تصادفی باشند و بر اساس قانون بیزین معادله به صورت زیر خواهد بود.

$$P(W|D, \propto, \beta, M) = \frac{P(D|W, \beta, M)P(W|\alpha M)}{P(D|\alpha, \beta, M)}$$ فرمول (۳)

در صورتی که فرض شود باقی مانده‌ها و وزن‌ها از توزیع گوسین پیروی کرده باشند، داریم:

$$P(D|W, \beta, M) = \frac{\exp(-\beta E_D)}{Z_D(\beta)}$$ فرمول (۴)

[1] Bayesian network

فرمول (۵) $$P(W|\alpha, M) = \frac{\exp(-\alpha E_W)}{Z_W(\alpha)}$$

برای اطمینان از این‌که $P(D|\alpha, \beta, M)$ به عنوان فاکتور تنظیم در معادله قرار گیرد معادله به صورت زیر خواهد بود:

فرمول (۶) $$P(W|D, \alpha, W\beta, M) = \frac{\exp(-F(W))}{Z_D(\alpha,\beta)}$$

آنگاه از مجموع معادلات خواهیم داشت:

فرمول (۷) $$P(D|\alpha, \beta, M) = \frac{Z_F(\alpha,\beta)}{Z_W(\alpha)Z_{D,}(\beta)}$$

که

فرمول (۸) $$Z_D(\beta) = \left(\frac{\pi}{\beta}\right)^{\frac{N}{2}}$$

فرمول (۹) $$Z_W(\propto) = \left(\frac{\pi}{\propto}\right)^{\frac{N}{2}}$$

فرمول (۱۰) $$Z_F(\propto, \beta) = (2\pi)^{\frac{N}{2}} \det^{\frac{-1}{2}}(H) \exp(-F(W^{MP}))$$

فرمول (۱۱) $$H = \beta\nabla^2 E_D + \propto \nabla^2 E_W$$

ماتریس Hessian تابع هدف F می‌باشد. در معادله روش‌های لگاریتم و مشتق‌گیری به ترتیب به کار گرفته شده است، با فرض این‌که معادله تفاضلی تابع برابر صفر قرار گیرد مقدار $P(\propto, \beta|D, M)$ به بیشینهٔ خود رسیده و احتمال اولیهٔ وزن‌ها مینیمم می‌شود. سپس α و β به صورت زیر بیان می‌شوند:

فرمول (۱۲) $$\propto^{MP} = \frac{\gamma}{2E_W W^{MP}}$$

فرمول (۱۳) $$\beta^{MP} = \frac{n-\gamma}{2E_D W^{MP}}$$

فرمول (۱۴) $$\gamma = N - \propto^{MP} \text{trace}^{-1} H^{MP}$$

که در آن n تعداد نمونه، N مقدار کل پارامترهای شبکه، γ تعداد پارامترهای مؤثر که نسبتاً اثر بیشتری در کاهش میزان خطای توابع دارند. در ابتدا α و β بر طبق الگوریتم لونبرگ مارکوانت فرض می‌شود و تابع هدف F(w) با آموزش شبکهٔ منظم بیزین به کمترین مقدار خود می‌رسد. α و β بر طبق معادلات قبل به روزرسانی شده و سپس مقادیر بهینهٔ توزیع اولیه به دست می‌آید و حداقل مقدار جدید F(w) محاسبه می‌گردد و در نهایت شبکه آنقدر تکرار می‌گردد تا به همگرایی برسد. [۱۵و۳۹]. در این پژوهش از نرم افزار GeNIe2 و GeneXpro Tools 4 استفاده شده است. مدل مفهومی شبکهٔ بیزین مشابه با برنامه‌ریزی بیان ژن است. لذا مقادیر ورودی به ترتیب X1، X2 جهت برآورد X3 می‌باشد. متغیرهای سنجش سلامت محیط در جدول (۲) نشان داده شده است.

**جدول ۲. متغیرهای سلامت محیطی**

| ردیف | نام شاخص | واحد متغیر |
|---|---|---|
| ۱ | رسوب | (میانگین رسوب تولیده شده – تن در سال) |
| ۲ | دبی | (متوسط دبی سالیانه به متر مکعب در ثانیه) |
| ۳ | دما | (متوسط دمای سالیانه به سلسیوس) |
| ۴ | تبخیر | (مجموع تبخیر سالیانه به میلیمتر) |
| ۵ | بارش | (متوسط بارندگی سالیانه به میلیمتر) |

فرآیند گام به گام حل یک مسئله با استفاده از برنامه‌ریزی بیان ژن متشکل از ۵ مرحله به شرح ذیل می‌باشد [۱۳ و ۴۸]: ۱- انتخاب مجموعه ترمینال که همان متغیرهای مستقل مسئله و متغیرهای وضعیت سامانه می‌باشد. مجموعه متغیرها محدودیتی نداشته و با توجه به اهداف پژوهش مشخص می‌شود. انتخاب تابع

برازش در این مرحله صورت می‌گیرد که معمولاً از جذر میانگین مربعات خطا استفاده می‌شود [۶۲، ۵۸]. ۲- انتخاب مجموعه توابع که شامل عملگرهای ریاضی، توابع آزمون و توابع بولی می‌باشد. عملگرهای ریاضی شامل ۱۰ عملگر ضرب، تقسیم، جمع، تفریق، جذر، لگاریتم، مجذور، مکعب و... می‌باشد که سه نوع جمع، تفریق و ضرب بیشترین استفاده را دارند اما در این پژوهش از کلیهٔ عملگرهای موجود در سیستم بیان ژن استفاده شده است. ۳- شاخص اندازه‌گیری دقت و صحت مدل که بر مبنای آن می‌توان مشخص کرد که توانایی مدل در حل یک مسئلهٔ خاص تا چه اندازه می‌باشد. ۴- مؤلفه‌های کنترل: مقادیر مؤلفه‌های عددی و متغیرهای کیفی که برای کنترل اجرای برنامه استفاده می‌شود. تعداد داده‌های بخش آموزش، تعداد داده‌های بخش آزمون، تعداد کروموزوم‌ها، اندازهٔ سر، تعداد ژن‌ها، انتخاب عملگر پیوند که با چهار گزینهٔ جمع، تفریق، ضرب و تقسیم قابل تنظیم هستند. ۵- معیارهای توقف برنامه: معیاری برای حصول نتیجه و توقف اجرای برنامه می‌باشد که مقادیر آن بر اساس میزان شکست در نتایج همبستگی با معیارهای بالاتر تولید نسل مبنا قرار گرفته است.

### ۵،۲. معیارهای ارزیابی کارایی مدل

در این پژوهش برای ارزیابی مدل‌های نامبرده از معیارهای زیر استفاده شد:

۱- جذر میانگین مربعات خطا (RMSE):

$$RMSE = \sqrt{\frac{\sum(Q_{obs}-Q_{pre})^2}{n}}$$ فرمول (۱۵)

تفاوت میان مقدار پیش‌بینی شده توسط مدل یا برآوردگر آماری و مقدار واقعی می‌باشد.

۲- ضریب تعیین ($R^2$):

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(Q_{obs}-Q_{pre})^2}{\sum_{i=1}^{N}(Q_{obs}-\bar{Q})^2}$$ فرمول (۱۶)

که در آن مقادیر $Q_{obs}$ و $Q_{pre}$ به ترتیب مقادیر مشاهده‌ای و مقادیر برآوردی می‌باشد. این معیار بیانگر اندازهٔ آماری نزدیک داده‌ها به خط رگرسیون برازش شده می‌باشد.

## ۳. نتایج

### ۱،۳. نتایج بررسی مدل برنامه‌ریزی بیان ژن

به منظور بررسی بهترین مقادیر تولید نسل در مدل GeneXpro Tools می‌توان به دو طریق عمل نمود. در بخش اول با انتخاب بهترین برازش می‌توان دقت بالایی را در مقادیر برآوردی به دست آورد اما این عمل با توجه به حجم داده‌های این پژوهش از ۸ ساعت تا ۷۲ ساعت به طول می‌انجامد، لذا در این پژوهش با بررسی تصادفی مقادیر تولید نسل، از مقادیر ۱۰۰۰ تا ۳۰۰ هزار بار استفاده شد و مجموعه عملگرهای موجود در سیستم در مورد مقادیر متغیرها اجرا شد و پس از بررسی مقادیر ضریب همبستگی در نقطه‌ای که مقدار تولید نسل کمترین تغییرات همبستگی را دارد به عنوان بهترین مقدار انتخاب شد. این مقادیر در جدول (۳) آورده شده است. مجموعه مدل به گونه‌ای طراحی شده است که داده‌های دو سال اخیر بیشترین همبستگی را با یکدیگر دارند، لذا داده‌های ورودی هر دو مدل برنامه‌ریزی بیان ژن و بیزین به صورت معادلهٔ زیر تعریف شده است. این رابطه نشان می‌دهد مقادیر هر متغیر در سال آینده به متغیرهای سال فعلی و سال گذشته مرتبط است و از لحاظ مفهومی دارای بیشترین خودهمبستگی می‌باشد. این مفهوم از نتایج پژوهش ابراهیمی و همکاران از طریق بررسی تحلیل عاملی و بررسی چند مقیاسی داده‌ها اثبات شده است [۹].

$$X3=X1+X2$$ فرمول (۱۷)

در این معادله $X_3$ متغیر آتی، $X_2$ متغیر فعلی، $X_1$ متغیر گذشته می‌باشد.

نتایج بررسی مدل برنامه‌ریزی بیان ژن در داده‌های آموزش سال ۱۳۶۹ تا ۱۳۸۴ و آزمون مدل در سال‌های ۱۳۸۵ تا ۱۳۹۲ در بخش دبی، رسوب، بارش، دما، تبخیر

با توجه مقادیر RMSE و $R^2$ مدل در جدول (۴) آورده شده است.

**جدول ۳. مقادیر تولید نسل و مقادیر همبستگی آن**

| ردیف | ۱۰۰۰ | ۵۰۰۰ | ۵۰۰۰۰ | ۱۰۰۰۰۰ | ۱۵۰۰۰۰ | ۲۰۰۰۰۰ | ۲۵۰۰۰۰ | ۳۰۰۰۰۰ |
|---|---|---|---|---|---|---|---|---|
| دبی | ۰/۵۲ | ۰/۶۱ | ۰/۷۳ | ۰/۷۹ | ۰/۸۲ | ۰/۸۷ | ۰/۸۸۵ | ۰/۸۸۵۶ |
| رسوب | ۰/۶۱ | ۰/۷۲ | ۰/۷۸ | ۰/۸۱ | ۰/۸۹ | ۰/۹۲ | ۰/۹۲۷ | ۰/۹۳ |
| دما | ۰/۵۶ | ۰/۶۸ | ۰/۷۴ | ۰/۸۶ | ۰/۸۸ | ۰/۹۱ | ۰/۹۲ | ۰/۹۲۷ |
| بارش | ۰/۲۶ | ۰/۳۷ | ۰/۵۳ | ۰/۶۴ | ۰/۷۵ | ۰/۸۹ | ۰/۹۰ | ۰/۹۰۹ |
| تبخیر | ۰/۲۸ | ۰/۳۶ | ۰/۴۹ | ۰/۵۸ | ۰/۶۱ | ۰/۷۷ | ۰/۷۹ | ۰/۸۲ |

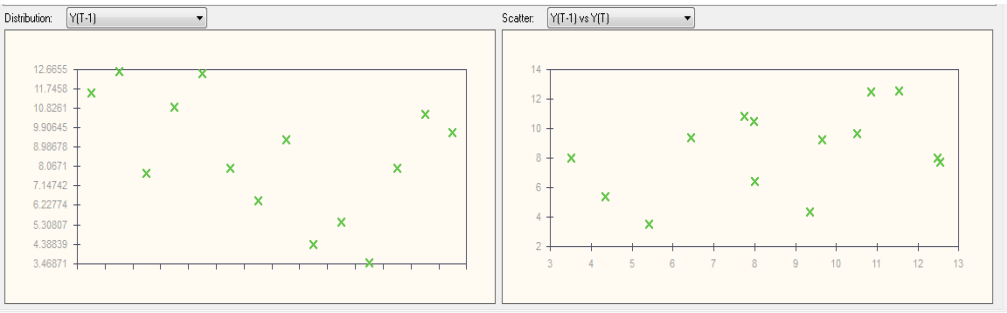
با توجه به مقدار فوق تولید نسل با مقدار ۲۰۰ هزار بار نسبت به مقادیر ۲۵۰ هزار و ۳۰۰ هزار شکست تغییرات اندکی دارد و بر این اساس به عنوان مبنای تحقیق انتخاب شد.

**جدول ۴. مقادیر ضریب تعیین و جذر میانگین مربعات خطای مدل برنامه‌ریزی بیان ژن**

| ردیف | عنوان متغیر | بخش آموزش | | بخش آزمون | |
|---|---|---|---|---|---|
| | | $R^۲$ | RMSE | $R^۲$ | RMSE |
| ۱ | دبی | ۰/۸۷ | ۱/۰۶ | ۰/۸۱ | ۱/۴۰ |
| ۲ | رسوب | ۰/۹۲ | ۱۱/۲۴ | ۰/۷۷ | ۳۹/۱۲ |
| ۳ | بارش | ۰/۸۹ | ۱۸/۲۴ | ۰/۸۱ | ۲۲/۱۴ |
| ۴ | دما | ۰/۹۱ | ۰/۱۵۰ | ۰/۷۹ | ۰/۱۹۸ |
| ۵ | تبخیر | ۰/۷۷ | ۸۵/۱۲ | ۰/۷۲ | ۹۳/۹۸ |

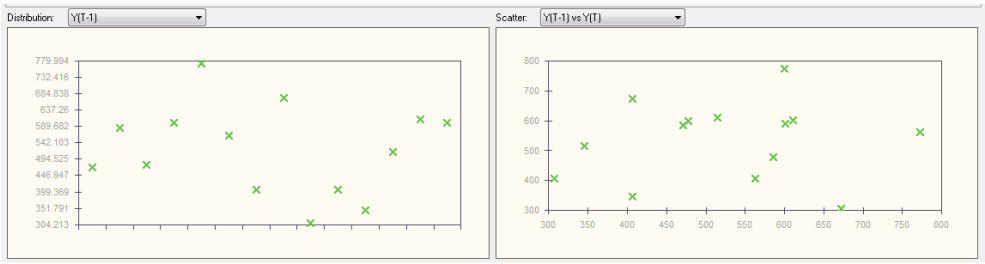
میانگین مقادیر ضریب تعیین برابر با ۰/۷۸ می‌باشد

در میان متغیرهای بررسی شده داده‌های متغیر رسوب بالاترین مقدار ضریب تعیین را در بخش آموزش و متغیر بارش و دبی بالاترین مقدار ضریب تعیین را در بخش آزمون دارا بودند. همچنین با توجه به نتایج ضریب تعیین در بخش آموزش مقادیر دما و رسوب انتظار می‌رفت نتایج بهتری را در بخش پیش‌بینی داشته باشد. در شکل ۲ تا ۶ چگونگی توزیع فضایی داده‌های مورد بررسی هر متغیر آورده شده است. بر اساس شکل‌های مورد بررسی ضریب تعیین کمتر مدل به دلیل پراکنش داده‌ها به صورت بسیار نامنظم و غیر خطی است.



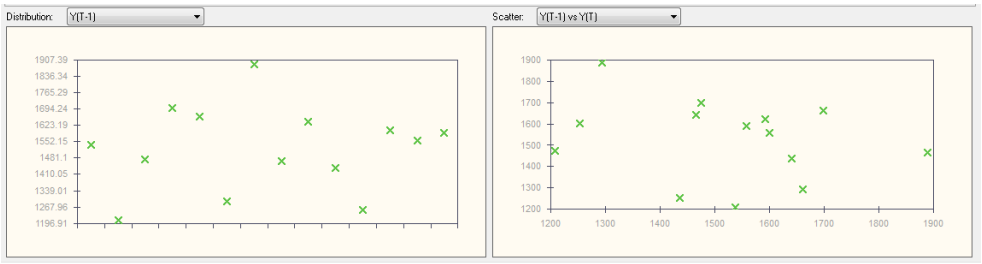

وضعیت پراکنش داده در بخش دبی حوزه آبخیز طالقان

**شکل ۲. توزیع فضایی داده‌های دبی**



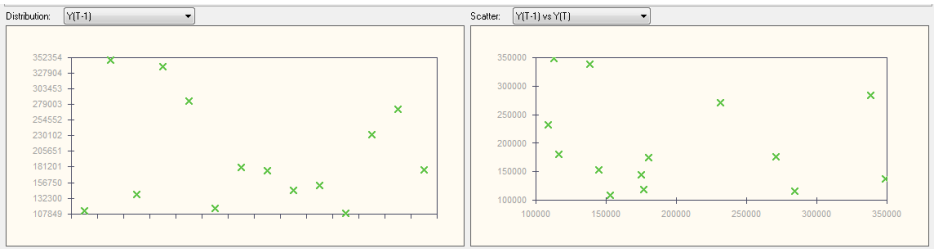

وضعیت پراکنش داده در بخش بارش حوزه آبخیز طالقان

**شکل ۳. توزیع فضایی داده‌های بارش**



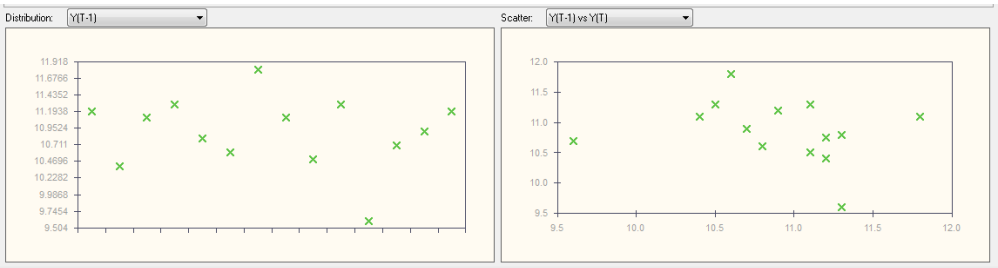

وضعیت پراکنش داده در بخش تبخیر حوزه آبخیز طالقان

**شکل ۴. توزیع فضایی داده‌های تبخیر**




وضعیت پراکنش داده در بخش مقادیر رسوب حوزه آبخیز طالقان

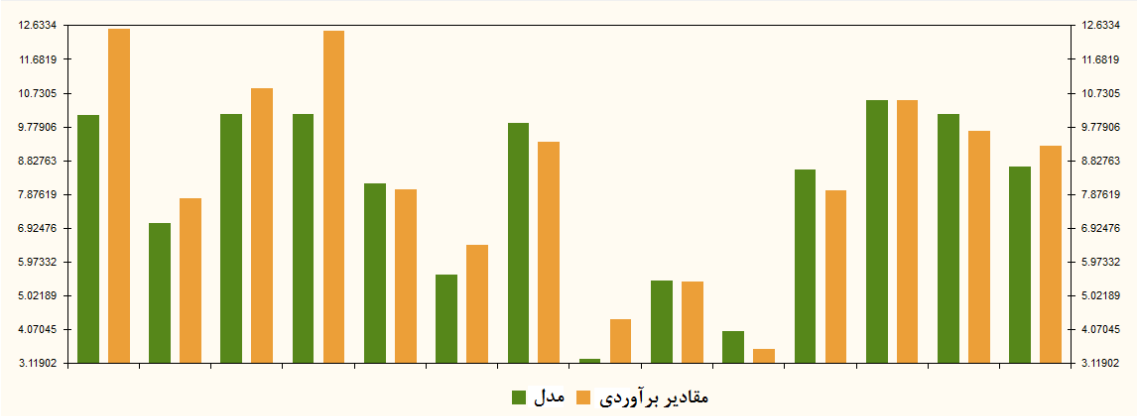
**شکل ۵. توزیع فضایی داده‌های رسوب**




وضعیت پراکنش داده در بخش دمای حوزه آبخیز طالقان

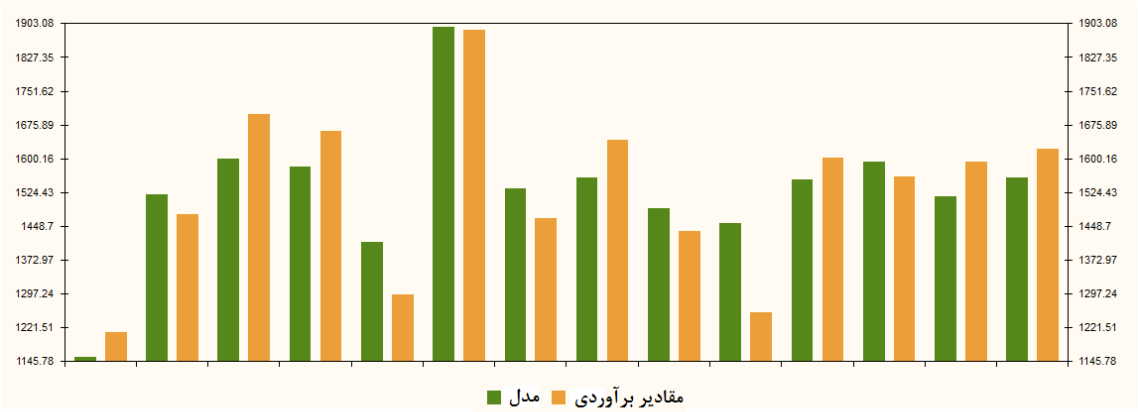
**شکل ۶. توزیع فضایی داده‌های دما**

به طور کلی به دلیل طول مدت آماری کمتر در بخش مدل دقت مدل نسبت به بخش آموزش کمتر است، زیرا مدل در بخش آموزش درک بهتری از فرآیند درونی داده ها دارد. شکل ۷ تا ۱۱ مقادیر پیش‌بینی مدل و مقادیر مشاهداتی را نشان می‌دهد.



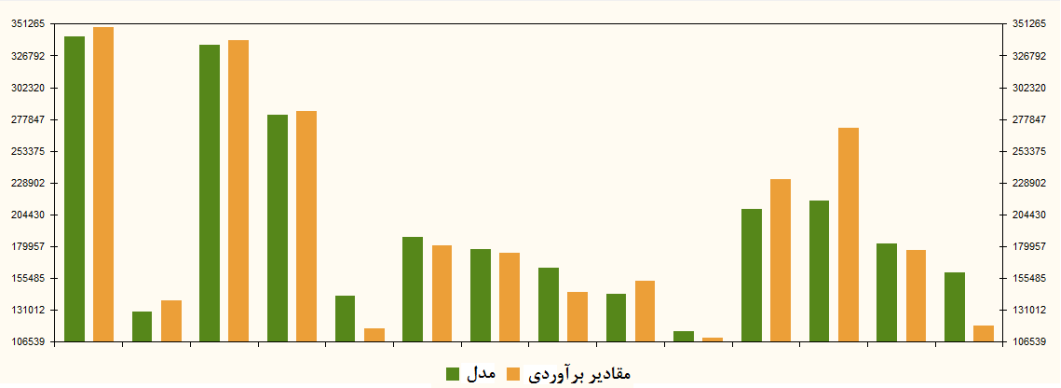

مقادیر پیش بینی مدل در بخش دبی

**شکل ۷. مقادیر پیش‌بینی مدل برنامه‌ریزی بیان ژن در بخش دبی**



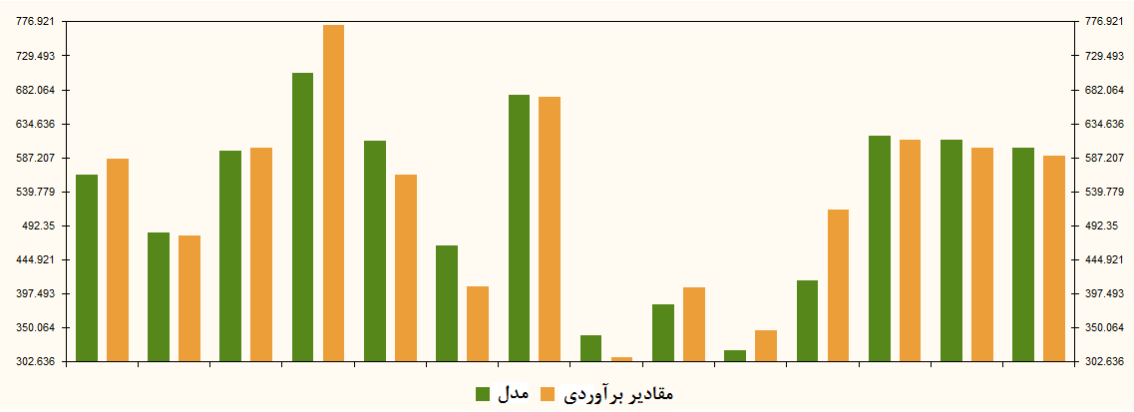

مقادیر پیش بینی مدل در بخش تبخیر

**شکل ۸. مقادیر پیش‌بینی مدل برنامه‌ریزی بیان ژن در بخش تبخیر**

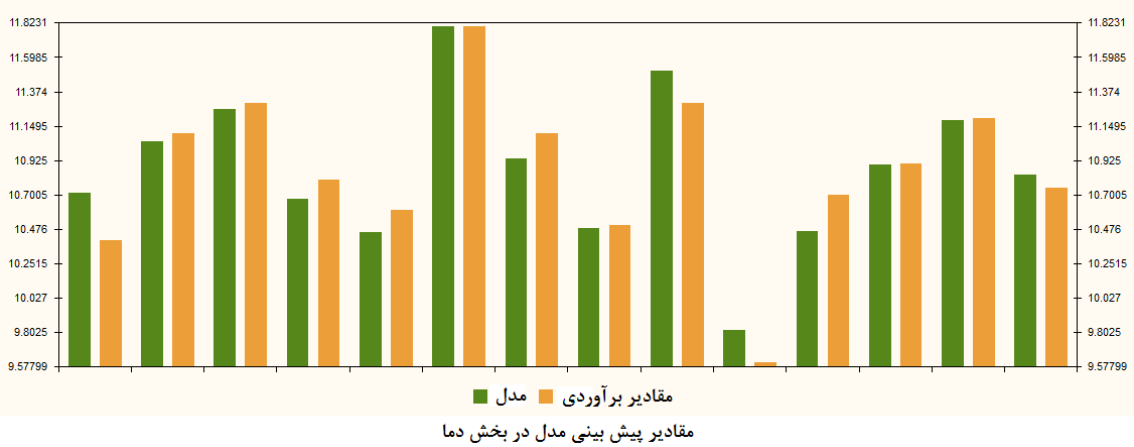



مقادیر پیش بینی مدل در بخش رسوب

**شکل ۹. مقادیر پیش‌بینی مدل برنامه‌ریزی بیان ژن در بخش رسوب**




**شکل ۱۰. مقادیر پیش‌بینی مدل برنامه‌ریزی بیان ژن در بخش بارش**




**شکل ۱۱. مقادیر پیش‌بینی مدل برنامه‌ریزی بیان ژن در بخش دما**

پس از آموزش و آزمون مدل مقدار سال ۱۳۹۵ کلیهٔ متغیرها به عنوان سال مورد پیش‌بینی انتخاب شد تا مقادیر سال ۱۳۹۳ و ۱۳۹۴ مبنای پیش‌بینی مدل با توجه به الگوی مفهومی طراحی شده مورد بررسی قرار گیرد. نتایج این بررسی در جدول (۵) آورده شده است. به طور کلی نتایج مدل در بخش رسوب، دبی، بارش مناسب و در بخش برآورد دما مطلوب نبوده است.

**جدول ۵. مقادیر دبی مشاهداتی و برآوردی در مدل برنامه‌ریزی بیان ژن در محدودهٔ مطالعاتی که بیانگر برآورد بیشتر در اکثر متغیرها دارد**

| ردیف | عنوان متغیر | مقدار مشاهداتی | مقدار برآوردی | تفاضل |
|---|---|---|---|---|
| ۱ | دبی | ۶/۶۷ | ۸/۵۱ | +۱/۸۴ |
| ۲ | رسوب | ۱۹۰۲۲۳ | ۱۹۶۵۴۷/۴ | +۶۳۲۴/۴ |
| ۳ | بارش | ۴۷۰/۸۵ | ۴۷۹/۶۲ | +۸/۷۷ |
| ۴ | دما | ۱۲/۱ | ۱۰/۶۰ | -۱/۵ |
| ۵ | تبخیر | ۱۸۰۲/۵ | ۱۵۸۸/۲۸۴ | -۲۱۴/۲۱۶ |

مقادیر جدول فوق نشـان می‌دهد در بخش متغیرهای دبی، رسـوب، بارش مدل بیان ژن دچار بیش برآوردی است و در مورد متغیر تبخیر و دما کم برآوردی حاصل و نتیجهٔ مدل اسـت. لذا در مورد پیش‌بینی سـال‌های آتی توسط مدل بیان ژن باید این تغییرات لحاظ شود تا دقت نتایج پیش‌بینی افزوده شود. مقادیر خطای مدل به صورت درصـدی از مقدار برآوردی در مقادیر سـال ۱۳۹۶ اعمال می‌شود. هر بخشی که دچار افزایش شده است با درصدی از خطا که افزایش یافته کاهش می‌یابد و در بخشـی که دچار افزایش برآورد مقادیر شـده، به صـورت درصـدی کاهش می‌یابد. نتایج مدل بیزین نیز نشـان داد در بخش دبی و بارش مقادیر مدل بیش از مقدار مشـاهداتی بوده و مقادیر رسـوب، تبخیر و دما کمتر از مقدار موجود در ایستگاه‌های سنجش بوده است. لذا مبنای پیش‌بینی مدل بیان ژن بوده که نتایج آن به تشریح آورده شده است. در مجموع عملکرد برنامه‌ریزی بیان ژن نشان داد دقت برآورد این مدل نسـبت به مدل بیزین بالاتر اسـت. مقیاس داده‌های هر بخش مرتبط با بخش دیگر نبوده و تأثیر گذار نیسـتند. از سـوی دیگر عملکرد نامناسـب مدل در بخش دما مربوط به مقیاس نیسـت، چرا که مدل ابتدا در مورد بیان ژن و نقشهٔ توالی انسان و مقادیر بسیار کوچک حضور ژن تا تعداد بسیار بالای مقادیر ژن طراحی شده است. از طرفی مطالعات در این زمینه کارایی مدل را اثبات می‌کند. نتایج ضـریب تعیین و جذر میانگین مربعات خطای مدل بیزین در جدول (۶) آورده شده است.

**جدول ۶. مقادیر دبی مشاهداتی و دبی برآوردی و مقادیر ضریب تعیین و جذر میانگین مربعات خطا در مدل بیزین**

| ردیف | عنوان متغیر | $R^2$ | RMSE | مقدار مشاهداتی | مقدار برآوردی | تفاضل |
|---|---|---|---|---|---|---|
| ۱ | دبی | ۰/۷۳ | ۱/۶۵ | ۶/۶۷ | ۹/۱۸ | +۲/۵۱ |
| ۲ | رسوب | ۰/۸۸ | ۱۵/۳۶ | ۱۹۰۲۲۳ | ۱۸۷۸۵۴/۴ | -۲۳۶۸/۶ |
| ۳ | بارش | ۰/۷۸ | ۲۲/۶۹ | ۴۷۰/۸۵ | ۴۹۱/۶۲ | +۲۰/۷۷ |
| ۴ | دما | ۰/۸۱ | ۰/۲۱۲ | ۱۲/۱ | ۱۰/۰۷ | -۲/۰۳ |
| ۵ | تبخیر | ۰/۷۴ | ۹۱/۰۸ | ۱۸۰۲/۵ | ۱۴۴۷/۱۱۲ | -۳۵۵/۳۸۸ |

میانگین مقادیر ضریب تعیین برابر با ۰/۷۸ می‌باشد

با توجه به توزیع نرمال داده‌ها در هر متغیر وضـعیت سـلامت حوزهٔ آبخیز طالقان به ۵ بخش تقسـیم شـد که نتایج آن در جدول (۷) آورده شده است. فرض بر آن است که داده‌های تمامی متغیرها از توزیع نرمال پیروی نموده و می‌توان براسـاس مقادیر کم تا زیاد در طول مدت آمار موجود به پنج رده تقسـیم‌بندی نمود. در این بخش مقادیر حداکثر و حداقل داده در طول دورهٔ آماری ۲۷ سـاله و نیز میانه داده‌ها مشخص شد. سپس بر مبنای مقادیر طبقات به صورت مساوی تقسیم شدند. مقادیر میان داده‌های هر طبقه مجدد میان‌یابی شـد تا حد طبقات بین داده‌ها نیز دسته‌بندی یکنواختی داشته باشند. مزیت این روش در آن است که هر حوزهٔ آبخیز براساس پتانسیل متغیرهای خود سـنجیده می‌شـود. مجموعه اعداد مورد اسـتفاده داده‌هایی اسـت که نرمال شـده و داده‌های پرت یا اوج از آن حذف شده تا توزیع یکنواختی در مورد داده صورت گیرد. مقادیر پیک فراوانی بسـیار اندکی دارند و تنش‌هایی که در حوزهٔ آبخیز تحت عنوان دبی حداکثر اتفاق می‌افتد وضعیت کلی را بیان نمی‌کند. مقادیر حدی زمانی اهمیت می‌یابند که هدف بررسی مقادیر حدی باشند و نه میانگین سالانهٔ دبی، رسـوب و یا تبخیر. به عنوان مثال اگر دمای حوزهٔ آبخیز طالقان در روز اول فروردین ۱۲/۱ باشد و در روز دوم ۱۳/۱ و در روز سـوم دوباره به مقادیر ۱۲/۱ نزدیک شـود، آن حوضـه سـلامتی خود را از دسـت داده اسـت. لذا میانگین تغییرات روند بهتری را از داده‌ها به نمایش می‌گذارد.

**جدول ۷. بررسی وضعیت سلامت متغیرهای مورد بررسی در حوزهٔ آبخیز**

| ردیف | وضعیت | نمره ارزشیابی | تبخیر | دما | بارش | رسوب | دبی |
|---|---|---|---|---|---|---|---|
| ۱ | سالم | ۵ | ۱۲۰۹ تا ۱۲۹۴ | ۹/۶ تا ۱۰/۴۷ | ۶۰۰/۶۲ یا بیشتر | ۱۳۴۹۳۱/۴۹ یا کمتر | ۱۰/۸۵ تا ۱۲/۵۴ |
| ۲ | بیمار | ۴ | ۱۲۹۴/۰۱ تا ۱۴۷۴/۵ | ۱۰/۴۸ تا ۱۰/۷۴ | ۵۷۴/۱۴ تا ۶۰۰/۶۱ | ۱۳۴۹۳۱/۵ تا ۱۵۲۸۲۷/۶۹ | ۹/۲۴ تا ۱۰/۸۴ |
| ۳ | سرطانی | ۳ | ۱۴۷۴/۶ تا ۱۶۲۰/۱ | ۱۰/۷۵ تا ۱۱/۲ | ۴۷۱/۲۱ تا ۵۷۴/۱۳ | ۱۵۲۸۲۷/۷ تا ۱۸۲۰۲۱/۲۹ | ۷/۷۷ تا ۹/۲۳ |
| ۴ | سرطان عود کننده | ۲ | ۱۶۲۰/۱۱ تا ۱۶۶۱ | ۱۱/۲۱ تا ۱۱/۳ | ۴۴۶/۸۸ تا ۴۷۱/۲۰ | ۱۸۲۰۲۱/۳تا ۲۵۶۰۵۷/۶۹ | ۶/۴۴ تا ۷/۷۶ |
| ۵ | سرطان پیشرفته | ۱ | ۱۶۶۱/۰۱ یا بیشتر | ۱۱/۳۱ یا بیشتر | ۴۴۶/۸۷ یا کمتر | ۲۵۶۰۵۷/۷ یا بیشتر | ۶/۴۳ یا کمتر |

جدول وضعیت سلامت حوزهٔ آبخیز در بخش محیطی [۹] مجموع امتیازات هر متغیر را در سال مورد بررسی به عنوان میزان سلامت حوزهٔ آبخیز معرفی می‌نماید. نتایج این پژوهش در جدول (۸) آورده شده است. با توجه به دقت بالاتر نتایج مدل برنامه‌ریزی بیان ژن، از داده‌های این بخش به عنوان خروجی وضعیت شاخص سلامت حوزهٔ آبخیز طالقان استفاده شد.

**جدول ۸. مجموع امتیازات برآورد شده در سال ۱۳۹۵ و وضعیت احتمالی حوزهٔ آبخیز طالقان در سال ۱۳۹۶**

| ردیف | وضعیت کیفی | وضعیت کمی | وضعیت (۱۳۹۵) | وضعیت (۱۳۹۶) |
|---|---|---|---|---|
| ۱ | سالم | ۲۴-۲۵ | - | - |
| ۲ | بیمار | ۲۰-۲۳ | - | - |
| ۳ | سرطانی | ۱۵-۱۹ | - | - |
| ۴ | سرطان عود کننده | ۱۰-۱۴ | - | * |
| ۵ | سرطان پیشرفته | ۹- ۵ | * (۸) | * |

با توجه به دقت مدل برنامه‌ریزی بیان ژن به میزان ۰/۸۷ پیش‌بینی می‌شود در سال ۱۳۹۶ نیز وضعیت حوزهٔ آبخیز طالقان در بهترین حالت سرطان عود کننده و در صورت حفظ وضعیت کنونی در وضعیت سرطان پیشرفته باقی خواهد ماند.

## ۴. بحث و نتیجه‌گیری

با توجه به نتایج این پژوهش وضعیت سلامت حوزهٔ آبخیز طالقان در بخش محیطی در بدترین شرایط ۲۷ سال اخیر سپری می‌شود. استفاده از مدل‌های داده کاوی می‌تواند وضعیت هر یک از متغیرها را بررسی و وابسته به مدت زمان هدف مورد مطالعه وضعیت پیش رو را پیش‌بینی نماید. پنج متغیر مورد بررسی در این پژوهش جزء مهم‌ترین متغیرهای ۵۰ سال اخیر در میان پژوهش‌های داده‌کاوی محیطی می‌باشد. از این رو انتخاب شاخص سلامت می‌تواند راه برنامه‌ریزی را براساس مقادیر این شاخص‌ها هموار نماید. این مدل بخشی از مدل بررسی سلامت و پایداری حوزهٔ آبخیز است که در بخش محیطی با دو مدل برنامه‌ریزی بیان ژن و باور بیزین به بررسی داده‌های مذکور جهت پیش‌بینی می‌پردازد و علت انتخاب این دو مدل بررسی نتایج ۲۷ پژوهش بین‌المللی است که توان این دو مدل را در بخش پیش‌بینی و الگو یابی داده ها به خوبی یاد می‌کند [۱۰]. از طرفی مقایسهٔ دو مدل برنامه‌ریزی بیان ژن و مدل بیزین نشان

داد الگویابی مدل برنامه‌ریزی بیان ژن دقت بالاتری داشته و می‌تواند مبنای پیش‌بینی قرار گیرد. نتایج این پژوهش مجموعه متغیرهای مربوط به محیط را مورد ارزیابی قرار داد که نتایج آن با نتایج عملکرد مدل به صورت تک متغییره در تحقیقات مختلف هم‌خوانی دارد [۲۹، ۶۱]. نتایج این پژوهش بیان می‌کند ضرایب تعیین مدل در بخش بیان ژن اگرچه با مقادیر برآورد شده متفاوت است اما این فاصله نسبت به مدل بیزین کمتر است. از سوی دیگر مدل بیان ژن در فرآیند آموزشی خود از عملگرهایی نظیر سینوس، کسینوس و آرک تانژانت که مجموعه داده‌های محیطی غالباً به این صورت هستند استفاده می‌نماید که نسبت به برازش خطی مدل بیزین کارا و دقت بالاتری را تولید می‌کند. میانگین مقادیر ضریب تعیین در مجموعه مدل بیان ژن ۰/۸۷ در مدل بیزن ۰/۷۸ می‌باشد. این تفاوت در الگوهایی که به صورت خطی با مقادیر مشاهده‌ای برازش داده شده‌اند می‌تواند دقت بالاتری داشته باشد. ارزیابی سلامت حوزهٔ آبخیز به صورت بومی تاکنون صورت نگرفته که نتایج این پژوهش می‌تواند مبنای ارزیابی سلامت سایر حوزه‌های آبخیز باشد. مقادیر مربوط به طبقه‌بندی متغیرها برای اولین بار با تأکید بر توان حوزهٔ آبخیز بررسی شده و مقادیر حدود طبقات یک محدوده به خوبی می‌تواند وضعیت قابل انتظار از حوزهٔ آبخیز را به تصویر بکشد. از این رو پیش‌بینی می‌شود تا با استفاده از داده‌های موجود در این زمینه شرایط ارزیابی سلامت و پایداری کل حوزه‌های آبخیز ایران با توجه به اهداف مد نظر صورت گیرد تا بتوان به‌منظور برنامه‌ریزی راهبردی جهت حل مشکلات و معضلات حوزه‌های آبخیز در ایران گام برداشت.

# References

[1] Ananda, J. and Herath, G. (2003). The use of analytic hierarchy process to incorporate stakeholder preferences into regional forest planning. Forest Policy Economics, 5:13-26.

[2] Berkes, F., Colding, J. and Folke, C. (2003). Navigating Social-ecological Systems: Building Resilience for Complexity and Change. Cambridge University Press, Cambridge, UK.

[3] Borji hasan gavyar, M., Moghaddamnia, A. and Sajedi, F. (2017). Evaluating the effectiveness of data-driven methods in predicting monthly rainfall. Journal of Earth Science Researches, 31: 42-61. (In Persian)

[4] Brick, P. (2001) Across the great divide: explorations in collaborative conservation and the American West. Island Press, Washington, DC.

[5] Burian, S.J., Brown, M.J. and McPherson, T.N. (2002). Evaluation of land use/land cover datasets for urban watershed modeling. Water Science Technology, 45(9): 269-276.

[6] Dastranj, A., Asadi Nalivan, O., Fallah, s., Saleh Nasab, A. and Jafari, Sh. (2015). Estimation of Annual Erosion and Sediment Using EPM and GIS Model (Case Study: Zidasht, Taleghan). Journal of Hydro Geomorphology, 1(4): 39-55. (In Persian)

[7] Dibyendu, S., Rupali, D., Avinandan, M. and Robyn, H. (2015). An Integrated Approach to Environmental Management. Wiley publisher.

[8] Dodds, W.K., Clements, W.H., Gido, K.R., Hilderbrand, H. and King, R.S. (2010). Thresholds, breakpoints, and nonlinearity in freshwaters as related to management. Journal of North American Benthological Society, 29: 988-997.

[9] Ebrahimi, P. (2018). Investigate and modeling of health and sustainability indices trend (case study: Taleghan watershed). PhD Thesis, University of Tehran.

[10] Finkelstein, J. and Jeong, I.C. (2017). Machine learning approaches to personalize early prediction of asthma exacerbations. Annals of the New York Academy of Sciences, 1387(1):153-65.

[11] Finlay, J., Hood, J.M., Limm, M.P., Power, M.E., Schade, J.D. and Welter, J.R. (2011). Light mediated thresholds in stream-water nutrient composition in a river network. Ecology, 92 (1), 140-150.

[12] Frair, J.L., Merrill, E.H., Beyer, H.L. and Maorales, J.M. (2008). Thresholds in land scape connectivity and mortality risks in response to growing road networks. Journal of Applied Ecology, 45(5): 1504-1513.

[13] Groselj, P. and ZadnikStirn, L. (2015). The environmental management problem of Pohorje, Slovenia: A new group approach within ANP-SWOT framework. Journal of Environment Management, 161:106-112.

[14] Hao, L.X., Sun, R.H. and Chen, L.D. (2014). Health assessment of river ecosystem in Haihe River Basin. China. Huan Jing Ke Xue, 35(10): 3692-3701.

[15] Hibbard, K., Jackson, S.T., Kubiszewski, I., Sinclair, P., Seorlin, S. and Steffen, W. (2011).Toward an integrated history to guide the future. Ecology and Society, 16 (4), 2.

[16] Izady, A., Abdalla, O., Sadeghi., M., Majidi, M., Karimi, A. and Chen, M. (2016). A novel approach to modeling wastewater evaporation based on dimensional analysis. Water Resources Management, 30(8): 2801-2814.

[17] Jansson, A.M., Hammer, M., Folke, C. and Costanza, R. (1994). Investing in Natural Capital: the Ecological Economics Approach to Sustainability. Washington, DC: Island Press.

[18] Kangas, A., Saarinen, N., Saarikoski, H., Leskinen, L.A., Hujala, T. and Tikkanen, J. (2010). Stakeholder perspectives about proper participation for regional forest programmes in Finland. Forest Policy and Economics, 12: 213-222.

[19] Karimi, S., Shiri, J., Kisi, O. and Shiri, A.A. (2015). Short-term and long-term stream flow prediction by using 'wavelet–gene expression' programming approach. ISH Journal of Hydraulic Engineering, 22(2): 148-162.

[20] Kavian, A., Bahrami, M. and Rohani, H. (2017). Comparison of Different Methods of Estimating Potential Evapotranspiration in River Flow Modeling Using SWAT Model in Taleghan Watershed. Journal of Desert Ecosystems Engineering, 6(15): 115-125. (In Persian)

[21] Kiker, G.A., Bridges, T.S., Varghese, A., Seager, T.P. and Linkov, I. (2005). Application of multicriteria decision analysis in environmental decision making. Integrated Environmental Assessment and Management, 1: 95-108.

[22] King, R.S. and Baker, M.E. (2010). Considerations for analyzing ecological community thresholds in response to anthropogenic environmental gradients. Journal of North American Benthological Society, 29: 998-1008.

[23] Kisi, O., Shiri, J. and Nazemi, A.H. (2011). A Wavelet-Genetic Programming Model for Predicting Short-Term and Long-Term Air Temperatures. Journal of Civil Engineering and Urbanism, 1(1): 25-37.

[24] Koundouri, P., Ker Rault, P., Pergamalis, V., Skianis, V. and Souliotis, I. (2016). Development of an integrated methodology for the sustainable environmental and socio-economic management of river ecosystems. Science of Total Environment, 540:90-100.

[25] Langarizadeh, M. and Moghbeli, F. (2016). Applying Naïve Bayesian Networks to Disease Prediction: a Systematic Review. Acta Information Medica, 24(5): 364–69.

[26] Langarizadeh, M., Khajehpour, E., Salari, R. and Khajehpour, H. (2017). Assessment of differential diagnosis of bacterial meningitis from other types of meningitis using fuzzy logic and neural networks. Payavard Salamat, 10(5):453-60. (In Persian)

[27] Lemos, M. and Rood, R. (2010). Climate projections and their impact on policy and practice. Wiley Interdisciplinary Review. Climate Change, 1(5): 670– 682.

[28] Li, Y., Li, Y. and Wu, W. (2016). Threshold and resilience management of coupled urbanization and water environmental system in the rapidly changing coastal region. Environmental Pollutant, 208: 87-95.

[29] Loghmanpour, N.A., Kormos, R.L., Kanwar, M.K., Teuteberg, J.J., Murali, S. and Antaki, J.F. (2016). A Bayesian Model to Predict Right Ventricular Failure Following Left Ventricular Assist Device Therapy. Heart Failure, 4(9):711-721.

[30] Mahan, C.G., Young, J.A., Miller, B.J. and Saunders, M.C. (2015). Using Ecological Indicators and a Decision Support System for Integrated Ecological Assessment at Two National Park Units in the Mid-Atlantic Region, USA. Environmental Management, 55(2): 508-522.

[31] Majidi Khalilabad, M., Alizadeh, A., Farid Hoseini, A. and Vazifedoust, M. (2017). Evaporation from Lakes and Reservoirs: Developing a Remote Sensing Algorithm of Reference and Water Surface Energy Balance. Journal of Iran-Water Resources Research, 13(2): 153-169. (In Persian)

[32] Majidi, M., Alizadeh, A., Farid, A. and Vazifedoust, M. (2015c). Lake and reservoir evaporation estimation: sensitivity analysis and ranking existing methods. Journal of Water and Soil, 29(2): 350-373 (In Persian)

[33] Majidi, M., Alizadeh, A., Farid, A. and Vazifedoust, M. (2016) Estimating evaporation from Doosti Dam Reservoir by development and application of a new method based on energy balance. Sharif Journal Civil Engineering. Articles in Press, Accepted Manuscript, Available Online from 19 March 2017. (In Persian).

[34] Majidi, M., Alizadeh, A., Farid. A. and Vazifedoust, M. (2014). Lake and reservoir evaporation: Energy balance estimations, evaluation of combination and radiation− temperature methods. Iranian Journal of Irrigation and Drainage, 8(3):603-616 (In Persian)

[35] Majidi, M., Alizadeh, A., Farid. A. and Vazifedoust, M. (2015a). Analysis of the effect of missing weather data in estimating daily reference evapotranspiration under different climatic conditions. Water Resources Management, 29: 2107-2124

[36] Majidi, M., Alizadeh, A., Farid. A. and Vazifedoust, M. (2015b). Estimating evaporation from lakes and reservoirs under limited data condition in a semi-Arid region. Water Resources Management, 29(10): 3711-3733.

[37] Marazza, D., Bandini, V. and Contin, A. (2010). Ranking environmental aspects in environmental management systems: a new method tested on local authorities. Environment International, 36(2):168-179.

[38] Martin, K.L. and Kirkman, L.K. (2009). Management of ecological thresholds to reestablish disturbance-maintained herbaceous wetlands of the south eastern USA. Journal of Applied Ecology, 46: 906-914.

[39] Martınez, A.M., Webb, G.I., Chen, S. and Zaidi, N.A. (2016). Scalable learning of Bayesian network classifiers. Journal of Machine Learning Research, 17(44):1-35.

[40] McMahon, T.A., Peel, M.C., Lowe, L., Srikanthan., R. and McVicar, T.R. (2013). Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis. Hydrology and Earth System Sciences, 17(4):1331-1345.

[41] Meissner, R. (2016). The Relevance of Social Theory in the Practice of Environmental Management. Science and Engineering Ethics, 22(5):1345-1360.

[42] Mendoza, G.A. and Prabhu, R. (2005). Combining participatory modeling and multicriteria analysis for community-based forest management. Ecology Management, 207: 145-156.

[43] Nicol, S., Fuller R.A., Iwamura, T. andChadès, I. (2015). Adapting environmental management to uncertain but inevitable change. Biological Sciences, 282: 1808-1829.

[44] Nordstrom, E.M., Eriksson, L.O. and Ohman, K. (2010). Integrating multiple criteria decision analysis in participatory forest planning: experience from a case study in northern Sweden. Forest Policy and Economics, 12: 562-574.

[45] Nordstrom, E.M., Romero, C., Eriksson, L.O. and Ohman, K. (2009). Aggregation of preferences in participatory forest planning with multiple criteria: an application to the urban forest in Lycksele, Sweden. Journal of Forest Researches, 39: 1979-1992.

[46] Nourani, V., Hosseini Baghanam, A., Adamowski, J. and Kisi, O. (2014). Applications of hybrid Wavelet-Artificial Intelligence models in hydrology, a review. Journal of Hydrology, 514: 358-377.

[47] Perotto, E., Canziani, R., Marchesi, R. andButelli, P. (2008). Environmental performance, indicators and measurement uncertainty in EMS context: a case study. Journal of Cleaner Production, 16:517–530.

[48] Petousis, P., Han, S.X., Aberle, D. and Bui, A.A. (2016). Prediction of lung cancer incidence on the low-dose computed tomography arm of the National Lung Screening Trial: A dynamic Bayesian network. Artificial Intelligence in Medicine, 72:42-55.

[49] Roboredo, D., Bergamasco, S.M.P.P. and Bleich, M.E. (2016). Aggregate index of social-environmental sustainability to evaluate the social-environmental quality in a watershed in the Southern Amazon. Ecological Indicators, 63: 337-345.

[50] Schaubroeck, T., Deckmyn, G., Giot, O., Campioli, M., Vanpoucke, C., Verheyen, K., Rugani,B., Achten, W., Verbeeck, H., Dewulf, J. andMuys, B. (2016). Environmental impact assessment and monetary ecosystem service valuation of an ecosystem under different future environmental change and management scenarios; a case study of a Scots pineforest. Journal of Environment Management, 173:79-94.

[51] Schensul, S.L., Nastasi, B.K. andVerma, R.K. (2006). Community-based research in India:a case example of international and trans disciplinary collaboration. Journal of Community Psychology, 38(1e2): 95-111.

[52] Sheppard, S.R. and Cizek, P. (2009). The ethics of Google Earth: crossing thresholds fromspatial data to landscape visualisation. Journal of Environment Management, 90 (6), 2102-2117.

[53] Shoaib, M., Shamseldin, A.Y., Melville, B.W. and Khan, M.M. (2015). Runoff Forecasting using HybridWavelet Gene Expression Programming (WGEP) Approach. Journal of Hydrology, 527: 326-344.

[54] Sima, S. and Tajrishi, M. (2014). Estimation of Urmia Lake evaporation using remote sensing data. Iran-Water Resources Research, 11(1):32-48. (In Persian)

[55] Sofonia, J.J. and Unsworth, R.K.F. (2010). Development of water quality thresholds during dredging for the protection of benthic primary producer habitats. Journal of Environment Monitoring,. 12 (1), 159-163.

[56] Taheri, A. (2004).Sedimentology of Taleghan River with special attention to sedimentation sensitivity of formations. Msc Thesis, Faculty of Geology, Tarbiat Moallem University. 110 pp. (In Persian)

[57] Tahmasebian, S., Ghazisaeedi, M., Langarizadeh, M., Mokhtaran, M., Mahdavi-Mazdeh, M. and Javadian, P. (2017). Applying data mining techniques to determine important parameters in chronic kidney disease and the relations of these parameters to each other. Journal of Renal Injury Prevention, 6(2): 83–7.

[58] Terzi, O. (2013). Daily pan evaporation estimation using gene expression programming and adaptive neural based fuzzy inference system, Journal Computing & Applications, 23(3): 1035-1044.

[59] Tress, G., Tress, B. and Fry, G. (2005). Clarifying integrative research concepts in landscape ecology. Landscape Ecology: 20 (4): 479-493.

[60] Utz, R.M., Hilderbrand, R.H. and Boward, D.M. (2009). Identifying regional differences in threshold responses of aquatic invertebrates to land cover gradients. Ecological Indicators, 9: 556-567.

[61] Xiong, G., Kola, D., Heo, R., Elmore, K., Cho, I. and Min, J.K. (2015). Myocardial perfusion analysis in cardiac computed tomography angiographic images at rest. Image Analysis, 24(1):77-89.

[62] Yassin, M., Alazba, A.A., Mattar, M.A. (2016). Artificial neural networks versus gene expression programming for estimating reference evapotranspiration in arid climate, Agricultural Water Management, 163: 110-124.

[63] Yuan, W. and Yang, K. (2011). A method of catchments health assessment under value-pressure model and its application in urbanized river network area: A case study in Shanghai, China. Chinese Geographical Science, 21(1): 102-109.

[64] Zhang, F., Zhang, J.Q., Wu, R., Ma, Q.Y. and Yang, J. (2016). Ecosystem health assessment based on DPSIRM framework and health distance model in Nansi Lake, China. Stochastic Environmental Research and Risk Assessment, 30(4): 1235-1247.