Archive of SID

یازدهمین همایش ملی علوم و مهندسی آبخیزداری ایران
توسعه مشارکتی در مدیریت حوزه‌های آبخیز
11th National Conference on Watershed Management Sciences and Engineering of Iran
Participatory Development in Watershed Management
۳۱ فروردین لغایت ۲ اردیبهشت ۱۳۹۵
April 19-21, 2016

تجزیه و تحلیل مولفه‌های اصلی موثر بر پیش بینی دبی جریان رودخانه

مجید خزایی[۱] ، عبدال شهریور[۲] و جعفر گوهرگانی[۳]

[۱]دانشجوی دکتری آبخیزداری دانشگاه هرمزگان

[۲]استادیار و عضو هیئت علمی مرکز تحقیقات آبخیزداری استان کهگیلویه و بویراحمد

[۳]کارشناسی بذر و نهال و سرپرست سازمان جهاد کشاورزی استان کهگیلویه و بویراحمد

چکیده

پیش بینی دبی جریان به عنوان یکی از مطرح ترین چالش های مدیریت منابع آب در دهه های اخیر بوده به طوری که محققان روش‌های متفاوتی را برای این امر در مقالات مختلف ارائه نموده و به کار گرفته اند. تمامی پدیده‌های هیدرولوژیکی به طور طبیعی پدیده های تصادفی هستند. این پدیده‌ها را می‌توان با استفاده از تعدادی متغیر تعیین کرد که ممکن است این متغیرها با یکدیگر دارای همبستگی زیادی باشند. مولفه‌های زیادی بر میزان دبی موثر است ولی اندازه‌گیری تمامی این مؤلفه‌ها در حوزه‌های آبخیز فراهم نمی‌باشد. از طرفی هدف از مدل سازی استفاده از روابط برای حوزه های با کمبود اطلاعات می باشد. بدین منظور از متغیرهای سهل الوصول بارش، دما و تبخیر که از متغیرهای اصلی موثر بر دبی نیز می باشند در تحقیق فعلی استفاده گردید. با توجه به اینکه مقادیر قبلی بارش، دما و رطوبت نیز از عوامل موثر بر تغییرات دبی می باشد از این عوامل نیز در پیش بینی دبی استفاده گردید. در مجموع ۹ متغیر انتخاب گردید و به کمک تجزیه و تحلیل مولفه های اصلی اقدام به تعیین سهم هر کدام از عوامل بر تغییرات دبی گردید. در نهایت با استفاده از رگرسیون اقدام به مدل سازی دبی گردید.

کلمات کلیدی: دبی، مدلسازی، مولفه های اصلی، رودخانه

مقدمه

پیش بینی دبی جریان به عنوان یکی از مطرح ترین چالش های مدیریت منابع آب در دهه های اخیر بوده به طوری که محققان روش‌های متفاوتی را برای این امر در مقالات مختلف ارائه نموده و به کار گرفته اند. تمامی پدیده‌های هیدرولوژیکی به طور طبیعی پدیده های تصادفی هستند. این پدیده‌ها را می‌توان با استفاده از تعدادی متغیر تعیین کرد که ممکن است این متغیرها با یکدیگر دارای همبستگی زیادی باشند (سیفی و همکاران، ۱۳۸۹).

بررسی مجموعه داده‌ها، ساده سازی و یافتن الگو ی حاکم بر متغیرها مهمترین مرحله از فرآیند مدل‌سازی و حل مساله است. در سال‌های اخیر روش‌های آماری چند متغیره در موضوعات مرتبط با منابع آب، هیدرو لوژی و محیط زیست به طور گسترده‌ای استفاده شده‌اند. روشهای چند متغیره رفتار همزمان چندین متغیر را مورد بررسی قرار می دهند و ساختار همبستگی متغیرها را درون گروه در نظر می گیرند (نیرومند، ۲۰۰۷). از آنجایی که اکثر متغیرهای تأثیرگذار در تمامی ایستگاه‌ها اندازه گیری نمی‌شود، بنابراین تحلیل داده‌های موجود برای درک اهمیت نسبی این متغیرها لازم است. یکی از روشهای آماری چندمتغیره که امروزه به طور وسیعی مورد استفاده قرار میگیرد، روش تحلیل عامل اصلی PFA و تحلیل مولفه اصلی PCA میباشد (ایونگ، ۲۰۰۵). روش مؤلفه‌های اصلی

www.SID.ir

Archive of SID

ابزاری ریاضی است که متغیرهای مؤثر را مشخص میکند. تحلیل مولفه اصلی به تبیین ساختار واریانس ،کوواریانس به کمک چند ترکیب خطی از متغیرهای اصلی سر و کار دارد. اهداف کلی آن عبارتند از :۱ کاهش حجم داده ها ؛ ۲: تعیین و تفسیر آنها.
اخیراً، استفاده از روش های PCA بر ای کاهش تعداد متغیرها ی ورودی و تفسیر بهتر نتایج به دست آمده از تحلیل داده‌های آب، رواج یافته است. یکی از کاربردهای مهم تحلیل مولفه های اصلی، در رگرسیون است. با استفاده از تحلیل مولفه های اصلی می توان تعداد زیادی متغیر توضیحی (متغیر مستقل) همبسته را با تعداد محدودی متغیر توضیحی جدید که مولفه های اصلی نامیده می شوند و ناهمبسته اند، جایگزین نمود. به این ترتیب نه تنها بعد مساله تقلیل می یابد بلکه مساله چند همخطی پیش نمی آید (لیو و همکاران، ۲۰۰۵).
این روش بین مجموعه‌ای بزرگ از داده‌های به ظاهر بی ارتباط، رابطه خاصی را تحت یک مدل فرضی برقرار میکند، سپس هر متغیر را به صورت ترکیب خطی از فاکتورهای پنهان تبدیل میکنددر این روش متغیرها به مجموعه های کوچکتری از عامل ها با کمترین اطلاعات از دست رفته و موثرترین اطلاعات موثر بر پدیده تبدیل میشوند .در تحلیل عاملی، متغیرهایی که در یک عامل تعریف می شوند، به یکدیگر کاملاً وابسته اند و این وابستگی، عامل را بوجود می‌آورد .از طرفی متغیرهای هر عامل هیچ وابستگی به متغیرهای عامل‌های دیگر ندارد (یزدان خواه، ۱۳۸۶).
خام چین مقدم و همکاران (۱۳۸۹) با پهنه بندی حداکثر بارش روزانه ایران از کلیه ایستگاه‌های بارانسنج کشورکه زیر نظر سازمان هواشناسی کشور و وزارت نیروست استفاده نمود. روش مؤلفه‌های اصلی برای حذف متغیرهای غیرضروری به کار گرفته شد که ضرورت فقط ۶ متغیرارتفاع، میانگین وانحراف معیار بارش سالانه و حداکثر روزانه و نسبت بارش زمستان به بهار از ۲۱ متغیر تأیید شد .
نوری و همکاران (۱۳۸۹) با پیش بینی ماهانه جریان با استفاده از ماشین بردار پشتیبان بر مبنای آنالیز مؤلفه اصلی با استفاده از ۱۸ متغیر ورودی به مدل، نشان داد که پیش پردازش متغیرهای ورودی به مدل با استفاده ازPCA بهبود عملکرد مدل را به همراه داشته است.
یو و یانگ (۱۹۹۶) روش PCA و آنالیز خوشه ای را برای داده های جریان روزانه ۳۴ ایستگاه جریان سنجی بکار بردند و نتایج آنها نشان داد که روش تجزیه خوشه ای سلسله مراتبی وارد و غیر سلسله مراتبی k-mean بطور قابل قبولی می تواند برای تفکیک مناطق همگن بکار رود و در نهایت سه ناحیه همگن برای منطقه بدست آوردند.
لی و چونگ (۲۰۰۹) از معادلات رگرسیونی برای تخمین دبی رودخانه استفاده نمودند. آنها از دو روش برای محاسبه دبی، یکی براساس میانگین سرعت و دیگری مبتنی بر منحنی سنجه در رودخانه مورد بررسی استفاده کردند. نتایج آنها نشان داد روش محاسبه دبی براساس سرعت دقت بالاتری را نشان داده است.

مواد و روش‌کار

PCA روش تقسیم یک ماتریس تشابه به یک مجموعه محورها یا مؤلفه‌های متعامد عمودی است که هر کدام از این محورها یک PC نامیده می‌شود به این روش مؤلفه‌ها را وزن دهی کرده و برای هر کدام یک مقدار ویژه بیان می‌کند (Pezhman و همکاران، ۲۰۰۹). مقدار ویژه هر محور برابر واریانس محاسبه‌شده برای آن محور است. درPCA مقادیر ویژه ماتریس تشابه بر اساس روند نزولی استخراج می‌شود به طوری که محورهای (مؤلفه‌های) متناظر PCA به طور متوالی مقادیر تغییرات را از بزرگ‌تر تا کوچک‌تر در ماتریس نشان می‌دهند. بنابراین نخستین محورهای PCA که واحدهای نمونه‌برداری روی آن‌ها مکان‌یابی خواهد شد،بیش‌ترین درصد مجموع تغییرات قابل تعریف را نشان خواهند داد (Moghadam و همکاران، ۲۰۰۳).

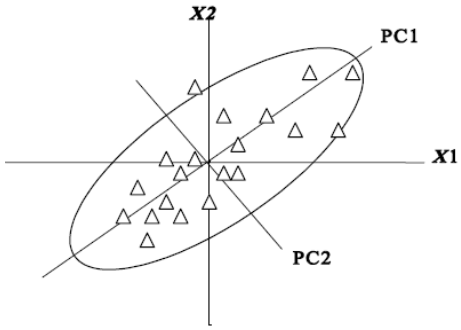
www.SID.ir

Archive of SID

اگر چه p متغیر برای مطالعه تغییرپذیری کل سیستم لازم است، ولی اغلب این تغییرپذیری را می‌توان با تعداد کمتر مثلاً k مؤلفه اصلی بیان نمود (k< p). در این صورت میزان اطلاعی که در k مؤلفه وجود دارد تقریباً مانند میزان اطلاع در p متغیر اولیه است. هر یک از k مؤلفه خود ترکیب خطی از چند متغیر اولیه است . بنابراین k مؤلفه اصلی را می‌توان به جای p متغیر اولیه به کار برد.

مجموعه داده‌های اولیه شامل p متغیر است که به مجموعه‌ای از داده‌های شامل k مؤلفه اصلی (ترکیب خطی متغیرهای مؤثر) کاهش داده می‌شود (چین مقدم و همکاران، ۱۳۸۹). مؤلفه‌های اصلی از نظر جبری ترکیبات خطی ویژه p متغیر تصادفی x1, x2 ,..., xp است .این ترکیبات خطی از نظر هندسی انتخاب یک دستگاه مختصات جدید را نشان می‌دهد که از دوران دستگاه اولیه با x1, x2 ,...,xp به عنوان محورهای مختصات به دست می‌آید. محورهای جدید جهت‌ها را با بیش‌ترین تغییرپذیری نشان می‌دهد. این جهت‌ها بردارهای ویژه ماتریس کوواریانس (Σ)یا ماتریس همبستگی(ρ) است که بیان ساده‌تری از ساختمان کوواریانس را فراهم می‌کند . مؤلفه‌های اصلی تنها به ماتریس کوواریانس(Σ) یا ماتریس همبستگی (ρ)مربوط به p متغیرx1, x2 ,..., xp بستگی دارد .مقادیر ویژه و بردارهای ویژه Σرا محاسبه و با این فرض که مجموع, واریانس متغیرها مساوی مجموع مقادیر ویژه است، مبادرت به تعیین سهم واریانس هر مؤلفه از کل می‌نماید .بردارهای ویژه جهت مؤلفه‌ها و مقادیر ویژه با اندازه آن‌ها متناسب است .اگر واحد اندازه‌گیری داده‌ها متفاوت باشد، طبق توصیه‌های انجام‌شده از ماتریس همبستگی(ρ) استفاده می‌شود (نیرومند، ۱۳۷۸ و Jollife، ۲۰۰۲).

شکل (۱) نقاطی را روی دو محور مختصات X_1 و X_2 نشان می‌دهد برای تعیین جهت عمومی نقاط، یک بیضی رسم می‌شود تا همبستگی بین متغیرها مشخص شود.



شکل ۱- انتقال داده‌ها به عوامل اساسی

برخی از نقاط خارج بیضی و البته تجمع تعداد زیادی از آن‌ها روی قطر اصلی بیضی مشاهده است جهت اصلی پراکنش نقاط نه در امتداد X1 و نه در امتداد X۲ بلکه بین آن‌ها و بیشتر در امتداد قطر اصلی بیضی است می‌شود .این محور PC1 نامیده می‌شود که اولین جزء اصلی تغییرپذیری X1 و X ۲ است. دومین جزء PC2 در امتداد قطر فرعی بیضی است که دقیقًا بر PC1 عمود بوده و باقی تغییرات در X ۱ و X۲ را شرح می‌دهد.۱ PC و PC۲ دو محور جدید برای شرح X ۱ و X۲ می‌باشند بنابراین می‌توان گفت X1و X۲ ترکیبی خطی از PC1 و PC۲ است یعنی:

$$Xi = aiPC1 + ai2PC2$$

همچنین می‌توان ارزش مؤلفه‌های اساسی را با استفاده از معادلات زیر به دست آورد:

www.SID.ir

Archive of SID

$$PC1 = w1X1 + w2X2$$

$$PC2 - w3X1 + w4X2$$

به طوری که wi ضریب رگرسیون اجزای اساسی روی متغیرها است.

مراحل تحلیل مؤلفه‌های اصلی

مرحله ۱: به دست آوردن اطلاعات

مرحله۲: کسر نمودن میانگین: برای این که آنالیز اجزای اصلی به طور صحیح کار کند باید میانگین را در هر بعد از داده‌ها کم کرد. کاستن میانگین یعنی میانگین سراسر هر بعد از مجموعه داده‌ها را کم کردن.

مرحله ۳:محاسبه ماتریس کوواریانس

مرحله ۴: محاسبه بردارهای مشخصه و مقادیر ویژه از ماتریس کوواریانس

مرحله ۵: متراکم سازی

مرحله ۶: وقتی که جزای اصلی را انتخاب شد (بردارهای مشخصه). داده‌ها حفظ و به شکل یک بردار ویژگی نشان داده می‌شوند و ترانهاده بردار محاسبه می شود.

مرحله ۷: مدلسازی رگرسیونی دبی با استفاده از مولفه های اصلی استخراجی

نتایج مدل‌سازی رگرسیونی

بعد از اینکه داده‌ها با انجام تبدیلات آماری نرمال گردید به منظور انجام مدل‌سازی وارد نرم‌افزارSPSS گردید. در ابتدا داده‌ها به دو دسته ۷۰ درصد و ۳۰ درصد جهت مدل‌سازی و اعتبارسنجی تقسیم گردیدند. تمام فرضیات مدل‌سازی رگرسیونی از جمله عدم هم خطی بین متغیرها و نرمال سازی باقیمانده‌ها روی‌داده‌ها آزمون گردید و سپس اقدام به انجام مراحل مدل‌سازی گردید. در ابتدا با استفاده از جدول آنالیز واریانس سطح معنی‌داری مدل‌ها مورد آزمون قرار می‌گیرد. در مرحله بعد پارامترهای مدل با توجه به سطح معنی‌داری آن‌ها انتخاب و معادله مدل‌ها استخراج گردید. به منظور بررسی دقت مدل ها اقدام به تعیین چندین اماره اعتبارسنجی از جمله ضریب تبیین و میزان ریشه خطا اقدام گردید در نهایت با استفاده از معادله‌های به دست آمده و تغییر ضرایب مدل ها اقدام به واسنجی مدل گردید.

نتایج

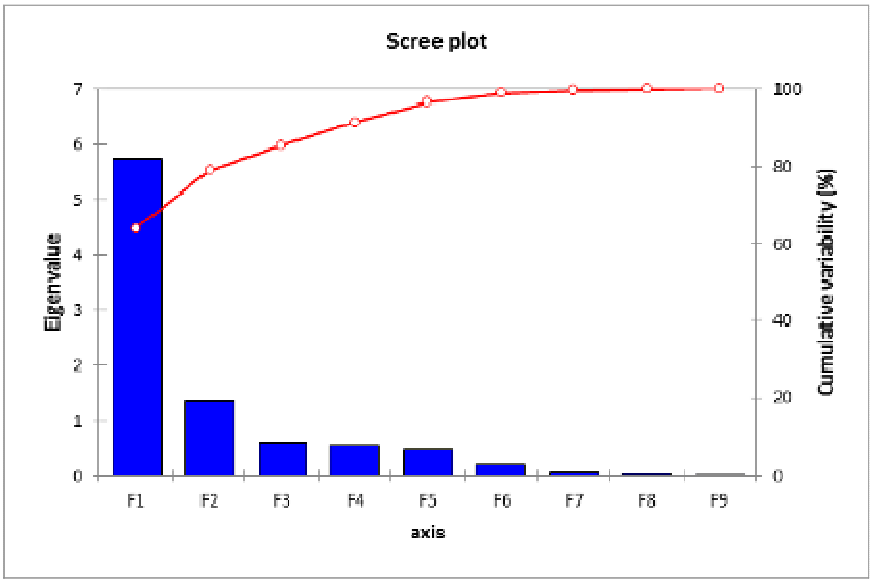
اولین گام در بررسی ارتباط بین متغیرها بررسی ماتریس همبستگی آن‌ها است. ابتدا مجموعه‌ای از متغیرها را انتخاب کرده و همبستگی بین آن ا را بررسی می‌کنیم .در صورت وجود همبستگی بین متغیرها، می‌توان عوامل را از روش مؤلفه‌های اساسی استخراج نمود.
گام اول در تجزیه و تحلیل مؤلفه‌های اصلی استخراج مؤلفه‌های اصلی می‌باشد که در این مطالعه ۱۰ عامل استخراج گردید که در جدول (۱) ارائه‌شده است. داده‌های جدول (۱) نشان داده است که ۷۶ درصد پراکندگی داده‌ها توسط دو مؤلفه اول بیان می‌شود. مقدار ویژه هر یک از مؤلفه‌های اصلی در شکل (۲) ارائه‌شده است که از این طریق می‌توان بهترین مؤلفه را مشخص نمود.

جدول۱- مقادیر ویژه و درصد تغییرپذیری بردارها

www.SID.ir

Archive of SID

عامل های اصلی	F1	F2	F3	F4	F5	F6	F7	F8	F9
مقدار ویژه	۵/۷	۱/۴	۰/۵۸	۰/۵۴	۰/۴۶	۰/۱۹	۰/۰۶	۰/۰۴	۰/۰۲
درصد تغییرپذیری(%)	۶۳/۸	۱۵	۶/۵	۶/۰۳	۵/۱۶	۲/۲	۰/۶۷	۰/۴۰	۰/۱۴



شکل۲-مقادیر ویژه بردارها (نمودار اسکری)

د مؤلفه اول بهترین انتخاب می‌باشد چنان که مشاهده می‌گردد مقدار ویژه اولین مؤلفه بزرگ‌تر از یک می‌باشد و حدود ۶۲ درصد پراکندگی مجموعه داده‌ها توسط این مؤلفه بازگو می‌شود. به عبارت دیگر همه معیارها نشان‌دهنده این است که انتخاب مؤلفه اول کافی است. سهم هر یک از متغیرها در واریانس تبیین شده به وسیله هر عامل در جدول (۲) ارائه شده است. بر این اساس مشاهده می‌شود سهم متغیرها در تبیین واریانس عامل اول نسبت به سایر ارقام مشابه مناسب‌تر می‌باشد.

جدول ۲- سهم متغیرها در تبیین واریانس عوامل

متغیرها	F1	F2	F3	F4	F5	F6	F7	F8	F9
بارندگی	۶/۶	**۱۵/۱**	۴/۶	۶۶/۲	۴/۷	۲/۵	۰/۰۱	۰/۰۸	۰/۰۰
بارندگی یک روز قبل	۸/۲	۰/۳	۸۳/۹	۳/۷	۲/۱	۱/۷	۰/۰۲	۰/۰۴	۰/۰۰
بارندگی دو روز قبل	۵/۵	۲۵/۸	۸/۲	۷/۵	۵۲/۲	۰/۴۲	۰/۳۲	۰/۰۰	۰/۰۰
درجه حرارت	۱۲/۶	۱۳/۵	۰/۱۴	۱/۹	۳/۴	۲۳/۲۳	۲۰/۳	۹/۹	۱۴/۹۳
درجه حرارت یک روز قبل	۱۶/۳	۰/۰۷	۰/۰۴	۰/۵۰	۰/۸۰	۲۴/۵	۰/۳۴	۱۷/۴	۴۲/۰۵

www.SID.ir

Archive of SID

درجه حرارت دو روز قبل	۱۱/۷	۱۶/۳	۰/۱	۰/۲۲	۱۲/۵	۱۱/۱۳	۳۲/۱	۲/۵	۱۳/۴۲
تبخیر و تعرق	۱۰/۸	۱۸/۵	۰/۹	۹/۱۵	۸/۵	۶/۸	۲۹/۵۶	۹/۳	۶/۵۰
تبخیر و تعرق یک روز قبل	۱۵/۵	۰/۴	۰/۹	۸/۴	۰/۳	۱۵/۳	۰/۵۰	۴۱/۱۸	۱۷/۵
تبخیر و تعرق دو روز قبل	۱۲/۵	۱۰/۱۵	۱/۱	۲/۴	۱۵/۳	۱۶/۴	۱۶/۸	۱۹/۵۶	۵/۶

بر اساس این جدول مشاهده می‌شود سهم متغیرها در تبیین واریانس عامل اول نسبت به سایر ارقام مشابه مناسب‌تر می‌باشد به عنوان مثال سهم هر یک از متغیرهای بارش، بارش روز قبل، بارش دو روز قبل، دما، دما ماه قبل و دما دو ماه قبل و تبخیر، تبخیر ماه قبل و تبخیر دو ماه قبل در عامل F1 به ترتیب ۶/۶۴، ۸/۲۱، ۵/۵۵، ۱۲/۶۵، ۱۶/۲۸، ۱۱/۷۶، ۱۰/۷۹، ۱۵/۵۶ و ۱۲/۵۵ می‌باشد. کسینوس مجذور متغیرها برای عامل‌های اصلی در جدول ۳ ارائه شده است.

جدول ۳-کسینوس مجذور متغیرها برای عامل‌های اصلی

متغیرها	F1	F2	F3
بارندگی	**۰/۳۸**	۰/۲۰	۰/۰۳
بارندگی یک روز قبل	۰/۴۷	۰/۰۰	**۰/۴۹**
بارندگی دو روز قبل	۰/۳۲	**۰/۳۵**	۰/۰۵
درجه حرارت	**۰/۷۳**	۰/۱۸	۰/۰۰
درجه حرارت یک روز قبل	**۰/۹۴**	۰/۰۰	۰/۰۰
درجه حرارت دو روز قبل	**۰/۶۸**	۰/۲۲	۰/۰۰
تبخیر و تعرق	**۰/۶۲**	۰/۲۵	۰/۰۱
تبخیر و تعرق یک روز قبل	**۰/۸۹**	۰/۰۱	۰/۰۱
تبخیر و تعرق دو روز قبل	**۰/۷۲**	۰/۱۴	۰/۰۱

نتایج جدول (۳) نشان می‌دهد که در مؤلفه اول از متغیرهای مورد بررسی ،مؤلفه بارش دو ماه قبل در فاکتور دوم و بارش یک ماه قبل در فاکتور سوم جای گرفته است ولی بقیه متغیرها در فاکتور اول قرار گرفته‌اند. که نشان‌دهنده این واقعیت است که متغیر بارش تا یک ماه و دو ماه قبل در اولویت کمتری در پیش‌بینی دبی برخوردارند.

نتایج مدل‌سازی رگرسیونی

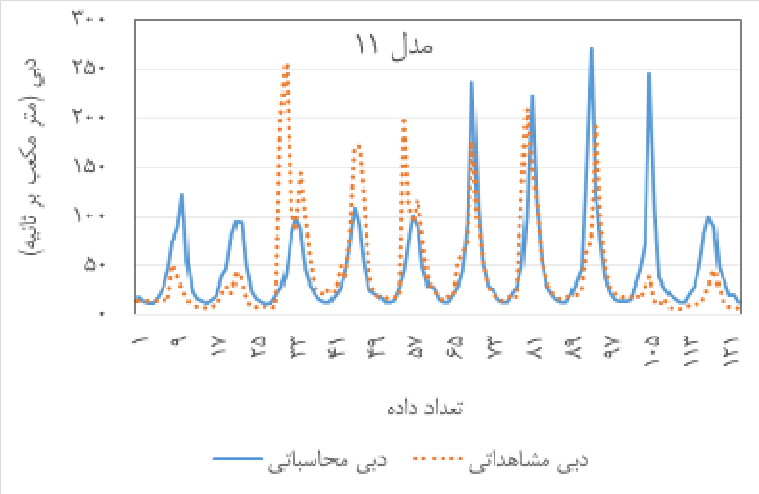
بعد از استخراج مدل‌های با ورودی مختلف اقدام به تعیین چندین اماره اعتبارسنجی از جمله ضریب تبیین و میزان ریشه خطا اقدام گردید و با استفاده از معادله‌های به دست آمده و تغییر ضرایب مدل ها اقدام به واسنجی مدل گردید. نتایج کلی مدل های شبیه سازی شده بعد از استخراج مولفه های اصلی (جدول ۴) و نمودار مقایسه‌ای بین مقادیر مشاهداتی و محاسباتی مدل برتر در شکل ۳ ارائه شده است.

www.SID.ir

Archive of SID

جدول ۴- نتایج اولویت‌بندی مدل‌ها با توجه به معیارهای ضریب تشخیص و میزان خطا

شماره	متغیر مستقل	R^2	RMSE	اولویت
۱	بارش	۰/۲۹	۰/۳۱	۹
۲	بارش یک ماه قبل	۰/۴۱	۰/۲۸	۶
۳	بارش دو ماه قبل	۰/۳۵	۰/۳۰	۷
۴	دما	۰/۱۸	۰/۳۲	۱۱
۵	دما یک ماه قبل	۰/۴۸	۰/۲۶	۴
۶	دما دو ماه قبل	۰/۶۱	۰/۲۲	۲
۷	تبخیر	۰/۱۷	۰/۳۲	۱۰
۸	تبخیر یک ماه قبل	۰/۴۳	۰/۲۷	۵
۹	تبخیر دو ماه قبل	۰/۵۴	۰/۲۴	۳
۱۰	بارش، دما و تبخیر	۰/۲۹	۰/۳۰	۸
۱۱	بارش، دما و تبخیر جاری، یک ماه و دو ماه قبل	۰/۶۵	۰/۲۱	۱



شکل ۳- تغییرات مقادیر مشاهداتی و محاسباتی دبی در دوره اعتبارسنجی در مدل ۱۱

www.SID.ir

Archive of SID

بحث و نتیجه گیری

با توجه به اینکه عوامل متعددی بر میزان دبی رودخانه‌ها تأثیر دارند و از طرفی بسیاری از عوامل موثر بر دبی به راحتی قابل دسترس و اندازه گیری نیستند در این بررسی از عوامل هواشناسی از جمله بارش، دما و تبخیر با تاخیرهای یک ماه و دو ما قبل جهت پیش‌بینی استفاده گردید. از طرفی این عوامل از مهم‌ترین عوامل موثر بر دبی می‌باشند و از سویی اندازه گیری و تهیه این داده‌ها دراکثر حوزه‌های آبخیزهای کشور و اکثر رودخانه‌ها موجود می‌باشند. به همین خاطر از ترکیبی از مؤلفه‌های بالا جهت پیش‌بینی دبی در این تحقیق استفاده گردید که در مجموع ۹ متغیر را شامل می‌گردند. که با ترکیب مدل های مختلف حدود ۱۱ مدل مورد شناسایی و تجزیه و تحلیل قرار گرفت. به منظور شناسایی مؤلفه‌های اساسی موثر بر دبی از بین مؤلفه‌های انتخابی با استفاده از روش تجزیه و تحلیل مؤلفه‌های اصلی اقدام به تعیین مهم‌ترین مؤلفه‌ها گردید. در روش PCA مؤلفه‌ها را وزن دهی کرده و برای هر کدام یک مقدار ویژه (Eigenvalue) بیان می‌کند، مقدار ویژه هر محور برابر واریانس محاسبه‌شده برای آن محور است. در نتیجه مؤلفه‌هایی که دارای بیش‌ترین وایانس باشند بیش‌ترین تأثیر را بر میزان دبی ایفا می کنند. در پژوهش حاضر ابتدا ماتریس همبستگی بین متغیرها ایجاد شد که تمامی متغیرها دارای همبستگی معنی‌داری بودند. نکته ای که در ماتریس همبستگی متغیرها مشاهده گردید همبستگی منفی درجه حرارت و تبخیر و تعرق و تاخیرهای یک ماه و دو ماه قبل و همبستگی مثبت بارش و تاخیر یک ماه و دو ماه قبل آن با هم بوده است. که نشان‌دهنده این موضوع است که افزایش دما هم سو با افزایش تبخیر و کاهش میزان بارش می‌باشد. در نتیجه رابطه معکوس بارش با سایر متغیرها قابل شهود بود. در مجموع ۱۰ مؤلفه اصلی استخراج شد که مقدار ویژه و درصد ورایانس و پراکندگی مؤلفه‌ها به ترتیب برای مؤلفه‌های یک تا ده معادل، ۶۳/۸۸، ۱۴/۹۹، ۶/۵۰، ۶/۰۳، ۵/۱۶، ۲/۲۰، ۰/۶۷، ۰/۴۱ و ۰/۱۴ بوده است. با توجه به نتایج به دست آمده از مؤلفه‌های به دست آمده می‌توان نشان داد که بیش‌ترین میزان تغییرات متغیرها توسط مؤلفه اول ایجاد شده است. ترکیب مؤلفه اول و دوم می‌تواند به عنوان مؤلفه‌های اساسی با میزان مقدار ویژه معادل ۷۸/۸۷ شناخته شود. با توجه به ضرایب محاسباتی متغیرها در بردار اول (F1) بیش‌ترین همبستگی را با مؤلفه اول به‌ترتیب متغیرهای درجه حرارت یک ماه قبل، تبخیر یک ماه قبل، دما، تبخیر دو ماه قبل، دما دو ماه قبل، تبخیر جاری و بارش دارا می باشند. هم چنین نتایج توزیع درصد سهم متغیرها در تبیین واریانس عامل اول نسبت به سایر ارقام مشابه مناسب‌تر می‌باشد به عنوان مثال سهم هر یک از متغیرهای بارش، بارش روز قبل، بارش دو روز قبل، دما، دما ماه قبل و دما دو ماه قبل و تبخیر، تبخیر ماه قبل و تبخیر دو ماه قبل در عامل F1 به ترتیب ۶/۶۴، ۸/۲۱، ۵/۵۵، ۱۲/۶۵، ۱۶/۲۸، ۱۱/۷۶، ۱۰/۷۹، ۱۵/۵۶ و ۱۲/۵۵ می‌باشد. به‌طور کلی نتایج تجزیه و تحلیل مؤلفه‌های اصلی نشان داد که همه متغیرهای مورد بررسی دارای اهمیت زیادی در تغییرات دبی هستند و همگی می‌توانند جزو مؤلفه اساسی مدل مورد بررسی قرار گیرند. بر اساس اهمیت متغیرها می‌توان گفت که در مؤلفه اول از متغیرهای مورد بررسی، مؤلفه بارش دو ماه قبل در فاکتور دوم و بارش یک ماه قبل در فاکتور سوم جای گرفته است ولی بقیه متغیرها در فاکتور اول قرار گرفته‌اند. که نشان‌دهنده این واقعیت است که متغیر بارش تا یک ماه و دو ماه قبل در اولویت کمتری در پیش‌بینی دبی برخوردارند. بعد از تعیین میزان اهمیت متغیرها و تعیین مؤلفه‌های اصلی اقدام به مدل‌سازی رگرسیونی گردید. همان طور که در فصل سوم ارائه شد به منظور مدل‌سازی رگرسیونی باید یک سری فرضیات را انجام داد تا نتایج به دست آمده دارای اعتبار بهتر و بالاتری گردند. در تجزیه و تحلیل رگرسیون از رگرسیون خطی دو و چند متغیره استفاده شد. در ابتدا با استفاده از آنالیز واریانس میزان معنی‌داری مدل مشخص گردید که از میان مدل های مورد بررسی تمامی مدل ها معنی‌دار گردیدند. سپس به کمک میزان سطح معنی‌داری ضرایب و پارامترهای مدل و معادله مدل استخراج گردید. با توجه به اینکه داده‌ها در ابتدا به دو بخش ۷۰ و ۳۰ درصد تقسیم گردید. مدل ها و روابط با استفاده از ۷۰ درصد داده‌های اولیه استخراج گردید سپس با استفاده از معادله‌های استخراج‌شده و ۳۰ درصد داده‌های کنار گذاشته شده میزان دبی محاسبه گردید. در نهایت مدل‌های با ضریب تشخیص بالاتر و مجذور میانگین مربعات خطای کمتر بعد از عمل واسنجی که با

www.SID.ir

Archive of SID

تغییر ضرایب معادله‌ها به‌صورت دستی صورت گرفت، انتخاب شد. هم‌چنین با ترسیم نمودار مشاهده‌ای و محاسباتی به بررسی مدل‌ها پرداخته شد. نتایج مدل‌سازی رگرسیونی نشان داد که اکثر مدل‌ها نتوانسته‌اند دبی را پیش‌بینی کنند. از میان مدل‌های مورد بررسی مدل شماره یازده با ورودی های بارشف دما و تبخیر و بارش و دما و تبخیر با یک ماه و دو ماه تبخیر دارای بیش‌ترین دقت و کمترین میزان خطا (R^2=۰/۶۵ و RMSE=۰/۲۰) می‌باشند. هم چنین این مدل توانسته است به خوبی مقادیر حداقل و حداکثر را شبیه‌سازی کند. مدل‌های شماره ۶ با ورودی دما دو ماه قبل و تبخیر دو ماه قبل به ترتیب در اولویت‌های دوم (R^2=۰/۶۱ RMSE=۰/۲۲) و سوم (R^2=۰/۵۴ RMSE=۰/۲۴) می‌باشند. هم چنین مدل‌های شماره ۴ و ۷ با ورودی‌های دما و تبخیر کمترین میزان دقت و بالاترین میزان خطا در اولویت آخر قرار گرفتند (R^2=۰/۱۸ RMSE=۰/۳۲؛ R^2=۰/۱۷ RMSE=۰/۳۲).

منابع

نوری، ر . فرخ نیا، ا .مرید، س . ریاحی مدوار، ح۱۳۸۹. تأثیر پیش‌پردازش متغیرهای ورودی به شبکه عصبی برای پیش‌بینی جریان . ماهانه با آنالیز مؤلفه‌های اصلی و موجک"،مجله آب و فاضلاب. ش . شصت و نه ص . ۱۳ تا ۲.

یزدان خواه س. ۱۳۸۷ .اهمیت نسب ی متغیرهای هواشناس ی در بر آورد تبخ یر-تعرق مرجع در اقل یم ه ای مختلف . پای ان نامه کارشناس ی ارشد آبیاری و زهکشی، دانشکده کشاورزی، دانشگاه تربیت مدرس، ۱۱۰ص.

خام چین مقدم فرهاد ، صدقی حسین ، کاوه فریدون ، منشوری محمد.۱۳۸۹. پهنه بندی حداکثر بارش روزانه ایران. نشریه آب و خاک (علوم و صنایع کشاورزی). جلد ۲۴ ، شماره ۱، .۱۳۸۹. ۹۷-۱۰۶.

Ouyang Y. 2005. A.plication of principal component and factor analysis to evaluate surface water quality monitoring network, Water Reservoir, 39: 2621-2635.

Liu C.W., Lin K.H. and Kuo Y.M. 2003. Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan, Science of the Total Environment, 313: 77- 89

www.SID.ir