بررسي كارايي روش‌هاي مختلف در بازسازي داده‌هاي هيدرومتري (بررسي موردي: حوزة آبخيز سفيدرود)

حجت‌ا... جليليان[١]، قباد رستمي‌زاد*[٢] و صالح آرخي[٣]
[١] دانش‌آموخته كارشناسي‌ارشد آبخيزداري، دانشگاه تهران، ايران
[٢] دانشجوي دكتري علوم آبخيزداري، دانشگاه تهران، ايران
[٣] استاديار دانشكده كشاورزي، دانشگاه ايلام، ايران
(تاريخ دريافت: ٨٩/١٠/٧، تاريخ تصويب: ٩٠/٨/٤)

چكيده

متاسفانه در ايران به دلايل مختلف از قبيل خرابي دستگاه‌ها، عدم برداشت داده‌ها، سوانح طبيعي و... بعضي داده‌ها در ايستگاه‌هاي هيدرومتري ثبت نشده‌اند كه موجب بروز نواقص آماري و ايجاد مشكلاتي در تجزيه و تحليل ايجاد مي‌كند. لذا ناگزير بايد اين داده‌هاي گمشده به طريقي برآورد شوند. روش‌هاي مختلفي براي بازسازي داده‌هاي گمشده وجود دارند كه هركدام با دخالت دادن متغيرهاي ويژه‌اي داده‌ها را برآورد مي‌كنند. در اين پژوهش از چهار روش رگرسيون تک متغيره، رگرسيون دو متغيره، لانگبين و توماس فيرينگ براي بازسازي داده‌هاي هيدرومتري شامل دبي ماهانه، دبي سالانه، رسوب و كيفيت آب بهره گرفته شده است. نتايج به‌دست آمده در اين پژوهش نشان داد كه بين روش‌هاي مختلف بازسازي دبي متوسط ماهانه در ١٧ ايستگاه مطالعاتي، از ١٩٢ مورد بازسازي، روش لانگبين با ٦٦ مورد (٣٣%) و روش رگرسيون دو متغيره با ٥٢ مورد (٢٧%) بهترين روش‌هاي بازسازي دبي متوسط ماهانه مي‌باشند، و روش رگرسيون دو متغيره از بين ١٧ ايستگاه در ٨ ايستگاه بهترين جواب را داده است و براي برآورد دبي متوسط سالانه مناسب مي‌باشد. همچنين نتايج اين پژوهش نشان داد كه از رابطه بين دبي- رسوب و دبي- كيفيت آب براي بازسازي اين داده‌ها با استفاده از روش‌هاي بكارگرفته شده در اين پژوهش نمي‌توان استفاده كرد.

واژه‌هاي كليدي: حوزه آبخيز سفيدرود، داده‌هاي هيدرومتري، روش لانگبين، روش رگرسيون دو متغيره

* نويسنده مسئول: تلفن: ٠٩١٨-٩٤٠٩٥٦٨ فاكس: ٩٨٢٦٣٢٢٤٩٣١٣+ E-mail: gh_rostamizad@ut.ac.ir

مقدمه

با توجه به اهميت داشتن سري آماري كامل و بدون نقص داده‌هاي هيدرولوژيكي مانند دبي جريان رودخانه‌ها و.....در مطالعات هيدرولوژي، برطرف كردن كمبودهاي آماري و تخمين داده‌هاي گمشده لازم و ضروري مي‌باشد. دست يافتن به يك روش مناسب كه بتواند تخمين درست و قابل اطميناني ارائه نمايد، مستلزم بررسي روش‌هاي مختلف پيشنهاد شده در اين زمينه مي‌باشد.

تمام بررسي‌هاي هيدرولوژي برپايه به‌دست آوردن سري‌هاي آماري مورد قبول پايه‌گذاري شده است، در صورتيكه اين اطلاعات ناقص باشد نمي‌توان يك تجزيه و تحليل آماري انجام داد. در مطالعات هيدرولوژي هر منطقه وجود اطلاعات هيدرومتري، پايه مطالعات را تشكيل مي‌دهد. اما متاسفانه بعضي مواقع اين اطلاعات مهم ناقص مي‌باشد و استفاده از آنها را با مشكل مواجه مي‌نمايد، به ويژه اگر در آن منطقه تعداد ايستگاه‌هاي هيدرومتري كم باشد، اين مشكل حادتر مي‌شود. بنابراين لازم است به روشي اين نواقص آماري بر طرف شود. زيرا به طور عموم نمي‌توان پروژه‌ها را به دليل نبود اطلاعات رها كرد. در بعضي ايستگاه‌هاي هيدرومتري حوزه¬آبخيز به علت خرابي دستگاه‌هاي اندازه‌گيري، عدم برداشت داده‌ها يا به علل گوناگون ديگر، بعضي اندازه‌گيري‌ها انجام نمي‌شود يا ثبت نمي‌گردد. از طرفي در محاسبات آماري داده‌ها لازم است كه در يك مجموعه تمامي ارقام و اعداد مربوط به يک كميت در يك دوره معين وجود داشته باشد لذا سعي بر اين است به جاي ارقام و اعداد گمشده از عددي استفاده شود كه نتيجه محاسبات را اشتباه ننمايد(Cadavid *et al.*, 1992). به دليل كوتاه مدت بودن آمارها در ايران، تكميل و استفاده از تمامي آمار موجود، لازم و ضروري است كه روش‌هاي مختلف بازسازي داده‌ها مورد دقت و ارزيابي قرار گيرد و بهترين روش برگزيده شود.

در خارج از كشور مطالعات موردي در زمينه كارايي روش هاي مختلف بازسازي داده هاي ايستگاه‌هاي هواشناسي، هيدرومتري و... صورت گرفته است كه به طور مختصر به آنها اشاره مي‌گردد. Domokos & (Kovac,1984) پژوهشي تحت عنوان تطويل سري‌هاي زماني دبي به وسيله رگرسيون را ارائه نمود. اين پژوهش در مورد دبي رودخانه زالا در مجارستان صورت گرفته است. در اين پژوهش با اعمال روش رگرسيون چند متغيره روي آمار ۲۵ ساله اين رودخانه مشاهده گرديد كه داده‌هاي برآورد شده و تطويل داده شده داراي اختلاف كمي با داده‌هاي واقعي مي‌باشند. (Clarke, 1979) از طريق برقراري روابط همبستگي بين بارش و دبي، مقادير دبي سالانه را بازسازي نمود و نتيجه گرفت كه از طريق روابط همبستگي با خطاي كم مي‌توان مقادير دبي سالانه را بازسازي نمود. (Sharma *et al.*, 1997)با استفاده از برآورد توابع چگالي، روشي براي شبيه‌سازي دبي جريان ارائه كردند كه آزمون آن توسط داده‌هاي ساختگي نشان داد كه اين روش در توليد وابستگي‌هاي غيرخطي بين مقادير متغير تصادفي، انعطاف پذيري بيشتري نسبت به روش‌هاي استوكاستيک دارد و از اين روش مي‌توان براي بازسازي داده‌هاي دبي استفاده نمود. (Hisdal & Tveito, 1993) با استفاده از توابع قائم تجربي اقدام به توليد داده‌هاي دبي ماهانه در ۹ حوزه در جنوب نروژ كرده است و سپس نتايج اين روش را با مدلهاي رگرسيون و مدلهاي مفهومي مقايسه كرده و به اين نتيجه رسيده‌اند كه روش توابع قائم تجربي جواب بهتري مي‌دهد. (Sen, 1978) با استفاده از مدل‌هاي رياضي از جمله مدل توماس فيرينگ نسبت به گسترش داده‌هاي دبي ماهانه اقدام كرده است و به اين نتيجه رسيده كه اين روش براي گسترش داده‌هاي دبي ماهانه مناسب مي‌باشد.

در داخل كشور نيز در زمينه بازسازي آمار هواشناسي و هيدرومتري مطالعات چندي صورت گرفته است.(Naghdi *et al.*, 2006) دو روش رگرسيون خطي چند متغيره و خودهمبستگي در بازسازي داده هاي دبي ماهانه در ايستگاه هيدرومتري اهواز با هم مقايسه كردند. نتايج بررسي اين روش‌ها با كمك دو آماره ميانگين قدرمطلق مربعات خطا[1] و درصد ميانگين انحرافات نسبي[1] نشان داد

[1] - Root Mean Squire of Error

که دقت روش رگرسيون چند متغيره نسبت به روش خودهمستگي بيشتر مي‌باشد. (Ashraf zadeh *et al.*, 2005) با استفاده از برآورد هسته‌اي چگالي نسبت به توليد آمار طولاني مدت براي حوزه رودخانه دز پرداختند و نتيجه گرفتند که اين روش براي توليد آمار طولاني مدت رودخانه‌ها مناسب مي‌باشد. (Sheikh, 1999) روش‌هاي مختلف بازسازي دبي حداکثر لحظه‌اي با استفاده از دبي متوسط روزانه در رودخانه‌هاي ايران را با هم مقايسه کرد و از بين روش‌هاي همبستگي، فولر، سنگال، لانگبين، اليس– گري، به اين نتيجه رسيد که در روش همبستگي، ٥٠ درصد برآوردها در محدوده ٢٠ درصد داده‌هاي مشاهداتي قرار داشتند و در ساير روش‌ها اين درصد کمتر از ٥٠ درصد بود.

به طور کلي هدف اين پژوهش تعيين بهترين روش گسترش داده‌هاي هيدرومتري در هر ايستگاه هيدرومتري موجود در منطقه مطالعاتي مي‌باشد. در ضمن فرضي‌هاي اين پژوهش عبارتند از:

١- روش‌هاي مختلف گسترش داده‌هاي هيدرومتري تفاوت معني‌داري با هم ندارند.

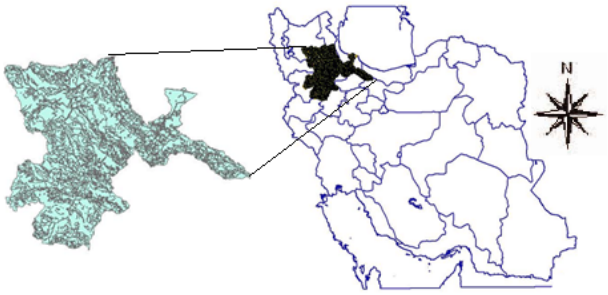
٢- روش توماس فيرينگ نسبت به ساير روش‌ها دقت بيشتري دارد.

٣- در هر ايستگاه هيدرومتري براي بازسازي آمار هيدرومتري، يك روش بهترين بازسازي را اعمال مي‌نمايد.

مواد و روش‌ها

منطقه مورد مطالعه

حوزة آبخيز سفيدرود بخشي از حوزة¬آبخيز درياي خزر محسوب مي‌شود. اين حوضه بين مختصات جغرافيايي ′٣٠,°٤٦ تا ′١٣,°٥١ طول شرقي و ′٥٥,°٣٤ تا ′٥٢,°٣٧ عرض شمالي قرار دارد. حوضه سفيدرود از شمال به حوضه تالش، مرداب انزلي، ناحيه ساحلي درياي خزر و حوضه ارس و از غرب به حوضه درياچه اروميه، از جنوب به حوزه خليج فارس و درياي عمان و از سمت شرق به حوضه مرکزي (زير حوزه درياچه نمك) محدود مي‌شود. مساحت حوزة آبخيز ٦٠٤٩٤ کيلومتر مربع است که حدود ٤٤٩٢٧ کيلومتر مربع آن را مناطق کوهستاني و ١٥٥٦٧ کيلومتر مربع را دشت‌ها و کوهپايه‌ها تشکيل مي‌دهد. دشت‌ها در بخش عليا و نيز بخش پاييني حوزه قرار داشته و مناطق مرکزي آن کوهستاني مي‌باشد. شهرهاي بيجار، زنجان، ميانه، هشترود، فيروزآباد، لوشان، منجيل، رشت، آستانه و لنگرود در اين حوزه قرار گرفته‌اند. شکل ١ موقعيت حوزه سفيد رود را در نقشه حوزه‌هاي ايران نشان مي دهد (sharifinejad, 1999).



شکل ١- موقعيت حوزه آبخيز سفيدرود

روش پژوهش

در اين حوضه چون بيشتر ايستگاه‌هاي هيدرومتري تازه تاسيس هستند يا اينکه داراي نواقص آماري خيلي زيادي مي‌باشند، در مجموع از بين ايستگاه‌هاي موجود در حوضه، تعداد ١٧ ايستگاه که داراي دوره آماري حداقل ٣٠ ساله بوده و در بالادست آنها سد نمي‌باشد، انتخاب شدند که مشخصات آنها در جدول ١ ذکر شده است.

[1] Mean Relative Deviation

جدول ۱- مشخصات ايستگاه‌هاي مورد استفاده در پژوهش

رديف	نام رودخانه	نام ايستگاه	طول جغرافيائي		عرض جغرافيائي		ارتفاع	سال آماري
۱	قزل اوزن	يساول (بيانلو)	۵۱	۴۷	۰۰	۳۶	۱۵۴۰	۸۴-۵۳
۲	تلوار	مهرآباد (سلامت آباد)	۵۱	۴۷	۴۰	۳۵	۱۶۵۰	۸۴-۴۰
۳	گاميشگايرود	هشتاد جفت	۵۳	۴۷	۱۰	۳۶	۱۴۵۰	۸۴-۵۳
۴	قزل اوزن	قره گوني	۵۷	۴۷	۱۴	۳۶	۱۴۱۰	۸۴-۵۴
۵	سوجاسرود	بنگي كند	۵۸	۴۷	۱۵	۳۶	۱۴۲۰	۸۴-۵۳
۶	تهم چاي	پالتي	۲۷	۴۸	۴۶	۳۶	۱۷۰۰	۸۲-۵۹
۷	زنجانرود	سرچم	۵۳	۴۷	۰۷	۳۷	۱۱۵۰	۸۴-۵۳
۸	قزل اوزن	پل دختر	۴۹	۴۷	۲۱	۳۷	۱۰۸۰	۸۴-۳۴
۹	آيدغموش	موتورخانه	۴۳	۴۷	۲۳	۳۷	۱۰۶۰	۸۴-۳۴
۱۰	شهرچاي	كوهسار	۲۳	۴۷	۳۱	۳۷	۱۵۰۰	۸۳-۵۱
۱۱	قزل اوزن	استور	۵۹	۴۷	۳۱	۳۷	۱۰۰۰	۸۴-۳۴
۱۲	قزل اوزن	گيلوان	۰۱	۴۹	۵۰	۳۶	۳۲۰	۸۳-۴۲
۱۳	شاهرود	گلينک (شاهرود)	۴۴	۵۰	۱۰	۳۶	۱۷۷۰	۸۳-۳۷
۱۴	الموت	باغكلايه	۲۷	۵۰	۲۲	۳۶	۱۳۵۰	۸۳-۵۳
۱۵	شاهرود	لوشان	۳۱	۴۹	۱۷	۳۶	۱۸۶۰	۸۳-۳۳
۱۶	توتكابن	توتكابن	۳۱	۴۹	۵۳	۳۶	۱۴۳۵	۸۳-۴۸
۱۷	شاري چاي	شاري چاي (ميانه)	۴۱	۴۷	۲۷	۳۷	۱۱۰۰	۸۴-۵۳

براي تهيه آمار و اطلاعات مورد نياز براي اجراي روش‌هاي گوناگون مورد بررسي، از آمار و اطلاعات موجود در بانک اطلاعات سازمان تحقيقات منابع آب ايران (تماب) استفاده و از بين متغيرهاي مختلفي كه در ايستگاه‌هاي هيدرومتري اندازه‌گيري شده است، آمار مربوط به دبي متوسط ماهانه، دبي متوسط سالانه، رسوب و كيفيت آب براي مطالعه انتخاب و در پايان براي ارزيابي و تجزيه و تحليل روش‌هاي مختلف از نرم افزارهاي Excel و SPSS استفاده گرديد.

با توجه به اين كه بازسازي به روش‌هاي مختلف در هر ايستگاه صورت مي‌گيريد؛ بنابراين بايد با استفاده ازيک شاخص از بين روش‌هاي مختلف يک روش را به عنوان بهترين روش گزينش نمود كه در اين پژوهش از شاخص اشتباه استاندارد براي تعيين بهترين روش استفاده شده است كه با استفاده از فرمول زير اين شاخص محاسبه مي‌شود. در ضمن براي كنترل كيفيت داده‌ها از روش ران تست استفاده شد كه داده‌ها از كيفيت مطلوب برخوردار بودند.

$$S_e = \sqrt{\frac{1}{n}\sum (Q_o - Q_c)^2} \qquad (1$$

كه در آن : S_e : اشتباه استاندارد n : تعداد داده‌ها

Q_o: مقادير مشاهداتي Q_c: مقادير برآوردي

با استفاده از اين آماره، روشي مناسب است كه داراي اشتباه استاندارد كمتري باشد.

براي اجراي روش‌هاي مختلف در هر ايستگاه از كل آمار هيدرومتري، ۲۰ درصد داده‌ها كنار گذاشته شد و فرض شد كه اين داده‌ها وجود ندارند (به عنوان داده

اعتبارسنجي)، سپس با روش‌هاي مختلف اقدام به بازسازي اين داده‌ها شد و سرانجام از طريق محاسبه اشتباه استاندارد بهترين روش در هر ايستگاه تعيين شد. لازم به يادآوري است كه در اين پژوهش براي از بين بردن تاثير منفي داده‌هاي پرت در برآوردها قبل از استفاده از هريک از روش‌ها اقدام به حذف داده‌هاي پرت شده است كه روش كار به صورت زير است.

برپايه پيشنهاد انجمن منابع آب آمريكا در سال ١٩٨١، چنانچه ضريب چولگي داده‌ها بيش از ٠/٤ باشد آزمون داده‌هاي پرت براي مقادير زياد بايد انجام گيرد. هنگامي كه ضريب چولگي داده‌ها كمتر از ٠/٤- باشد آزمون براي داده‌هاي كم صورت مي‌گيرد و چنانچه ضريب چولگي بين ٠/٤ و ٠/٤- باشد آزمون بايد براي داده‌هاي زياد و كم انجام شود. براي آزمون داده‌هاي پرت براي مقدار زياد و يا كم به ترتيب مي توان از رابطه‌هاي زير استفاده نمود (Mahdavi, 2004):

$$y_H = \bar{y} + K_N S_y \qquad y_L = \bar{y} - K_N S_y \qquad (٢$$

كه در آنها : y_H و y_L، لگاريتم اعشاري آستانه‌هاي بالا و پائين داده‌هاي پرت، : $\bar{y}$ ميانگين لگاريتم داده-ها، K_N، ضريب روش داده‌هاي پرت كه از جدول مقادير K در سطح اطمينان ١٠ درصد به‌دست مي‌آيد، S_y، انحراف معيار لگاريتم داده‌ها مي‌باشد.

چنانچه در بين داده‌هاي موجود، داده‌هايي وجود داشته باشند كه از y_H بيشتر و يا از y_L كمتر باشند، جزء داده‌هاي پرت هستند. دبي‌هاي اوج كه جزئ داده‌هاي پرت قرار مي‌گيرند بايد با داده‌هاي تاريخي و يا داده‌هاي حوضه‌هاي مجاور مقايسه شوند. داده‌هاي پرت كمتر از آستانه پائيني، در دبي‌هاي حداكثر بايد از مجموعه داده‌ها حذف و سپس برپايه داده‌هاي باقي مانده، توزيع مناسب انتخاب و اين آزمون براي داده‌هاي موجود انجام گيرد كه در اين مطالعه در سطح ١٠ درصد داده پرت وجود نداشت(Mahdavi, 2004).

امروزه روش‌هاي مختلفي براي بازسازي و تكميل داده‌هاي هيدرومتري موجود مي‌باشد كه بسته به شرايط هر منطقه مي‌توان از آنها بهره جست. در زير چهار روش مورد استفاده در اين پژوهش شرح داده شده‌اند:

روش لانگبين

اين روش براي حالتي است كه ايستگاه ناقص داراي دوره آماري كوتاه و ايستگاه مجاور آن داراي دوره آماري طولاني باشد مناسب است. در اين پژوهش با كمک ماتريس همبستگي بين ايستگاه‌ها، ايستگاهي كه بالاترين ضريب همبستگي را با ايستگاه ناقص دارد انتخاب شدند. هر دو ايستگاه بايد در يك حوزه آبخيز واقع شده باشند و به اين ترتيب مي‌توان مقادير دبي را در ايستگاه ناقص به كمک معادله زير بازسازي كرد (Alizadeh, 2003):

$$y = a + bx \qquad (٣$$

كه در آن :y : مقادير دبي بازسازي شده براي ايستگاه ناقص كه به صورت لگاريتمي مي‌باشد، a :ميانگين لگاريتمي داده‌هاي ايستگاه ناقص، x : مقادير دبي لگاريتمي براي ايستگاه كامل در سالهائي كه ايستگاه ناقص داراي نواقص آماري مي‌باشد و b : از معادله زير به‌دست مي‌آيد.

$$b = \frac{\sum xy_i}{\sum x_i^{\,2}} \qquad (٤$$

كه در آن: x_i از تفاضل مقادير دبي لگاريتمي از ميانگين لگاريتمي اعداد دبي براي ايستگاه كامل در سالهائي كه هر دو ايستگاه داراي آمار مي‌باشند به‌دست مي‌آيد، y_i: از تفاضل مقادير دبي لگاريتمي ايستگاه كامل از ميانگين لگاريتمي اعداد دبي ايستگاه ناقص در سالهائي كه هر دو ايستگاه داراي آمار مي‌باشند به‌دست مي‌آيد و ضريب همبستگي r به‌صورت زير محاسبه مي‌شود.

$$r = \frac{\sum xy}{\sum x^2 \sum y^2} \qquad (٥$$

و با آنتي لوگ گرفتن از مقادير y مقادير دبي براي ايستگاه داراي نواقص آماري به‌دست مي‌آيد. لازم به يادآوري است از روش لانگبين براي گسترش داده‌هاي دبي متوسط ماهانه و سالانه استفاده شده است.

روش توماس فيرينگ

اين روش مي‌تواند براي بازسازي مقدار دبي ماهانه به‌كار مي‌رود. اين مدل يك دنباله از زنجيره ماركوف است كه فرض مي‌كند مقادير دبي ماهانه فقط به دو متغير شامل ميانگين و انحراف معيار مقادير همان ايستگاه وابسته است. در اين روش هر داده دبي ماهانه فقط به داده دبي ماه قبل از خود وابسته است. در اين روش از رابطه (۶) استفاده مي‌شود.

$$q_{j+1} = q_{avj+1} + b_{j,j+1}\left(q_j - q_{avj}\right) + z_i s_{j+1}\sqrt{1-r_{j,j+1}^2} \qquad (٦$$

كه در اين معادله: q_j, q_{j+1}: مقادير دبي ماهانه در ماه‌هاي (j) و (j+1) ، q_{avj}, q_{avj+1}: ميانگين مقادير دبي ماهانه براي ماه‌هاي (j) و (j+1)، s_j, s_{j+1}: انحراف معيار براي ماه‌هاي (j) و (j+1)، $r_{j,j+1}$: ضريب همبستگي بين مقادير دبي ماه‌هاي (j) و (j+1) و z_i: عدد تصادفي است كه بين صفر تا ۱ متغير بوده و از جدول اعداد تصادفي استخراج مي‌شود.

$$b_{j,j+1} : \left(s_{j+1} / s_j\right) r_{j,j+1} \qquad (٧$$

اگر j را ماه ويژه در نظر بگيريم، j+1 ماه بعدي آن مي‌باشد. لازم به ذكر است از روش توماس فيرينگ فقط براي گسترش داده‌هاي دبي متوسط ماهانه استفاده شده است.

روش رگرسيون تك متغيره

در بازسازي به كمك اين روش، مراحل زير به ترتيب انجام شده است:

- تعيين ماتريس همبستگي هر ايستگاه و ايستگاهي كه بالاترين ضريب همبستگي را با آن ايستگاه داشته است، براي بازسازي آمار ايستگاه مورد نظر انتخاب شده است.
- سپس معني‌داري مقدار ضريب همبستگي بين ايستگاه مورد نظر با ايستگاهي كه با آن بالاترين ضريب همبستگي را داشته است در سطح ۱ و ۵ درصد بررسي مي‌شود (تست فيشر) و در صورت معني‌دار بودن استفاده مي‌شود.
- معادله رگرسيون بين ايستگاه پايه به‌عنوان متغير وابسته و ايستگاهي كه بالاترين ضريب همبستگي را با آن دارد به عنوان متغير مستقل برقرار مي‌شود.
- در پايان به‌كمك معادله مذكور مقدار دبي براي ايستگاه مورد نظر بازسازي مي‌شود.

براي اطمينان از اين كه رگرسيون خطي برازش مناسبي را ارائه نموده است، اقدام به بررسي ديگر مدلهاي لگاريتمي، نمائي، دوجمله‌اي و تواني روش رگرسيون گرديد.

روش رگرسيون دو متغيره

در اين روش يک ايستگاه به عنوان متغير وابسته و دو ايستگاهي كه بالاترين ضريب همبستگي را با ايستگاه وابسته دارند به عنوان متغير مستقل در نظر گرفته مي‌شوند و معادله بين اين ايستگاه‌ها محاسبه مي‌شود. روش بهره‌گيري شده در رگرسيون دو متغيره، روش رگرسيون گام به گام مي‌باشد. در اين روش متغيرهاي مستقل يكي يكي وارد مدل مي‌شوند (گام به گام رو به جلو) و در صورت داشتن سهم ناچيز يک به يک از مدل حذف مي‌شوند. بنابراين فقط متغيرهائي در مدل مي‌مانند كه وابستگي مناسب را با متغير وابسته داشته و ضريب همبستگي آن معني‌دار باشد(Patra, 2001).

براي بازسازي داده‌هاي رسوب تنها از روش رگرسيون استفاده شده است(به دليل اينكه داراي اشتباه استاندارد

بيشتري بوده و مقادير برآوردي با مقادير مشاهداتي تفاوت زيادي دارند) و سپس مقادير اشتباه استاندارد اين مقادير برآوردي در ايستگاه‌هاي مختلف محاسبه شده است.

براي بازسازي داده‌هاي كيفيت آب نيز تنها از روش رگرسيون خطي استفاده شده است (به دليل اينكه داراي اشتباه استاندارد بيشتري بوده و مقادير برآوردي با مقادير مشاهداتي تفاوت زيادي دارند). در اين مطالعه، از رابطه همبستگي بين مقادير دبي و كيفيت آب يک ايستگاه استفاده شده است. همچنين، آمار مربوط به هدايت الكتريكي، كربنات، كلر، سولفات، مجموع آنيون، كلسيم، منيزيم، پتاسيم، مجموع كاتيون، درصد سديم، نسبت جذب سديم بررسي شد و از رابطه همبستگي بين دبي ثبت شده در ايستگاه با هر كدام از اين آمارها به‌صورت جداگانه استفاده شده است. سپس مقادير اشتباه استاندارد براي هر يک از پارامترهاي كيفيت آب در ايستگاه‌هاي مختلف محاسبه شد. براي برآورد مقادير هدايت الكتريكي از رابطه رگرسيوني بين مقادير هدايت الكتريكي با مقادير دبي متناظر با آن در هر ايستگاه استفاده و سپس مقدار آماره اشتباه استاندارد محاسبه گرديده است.

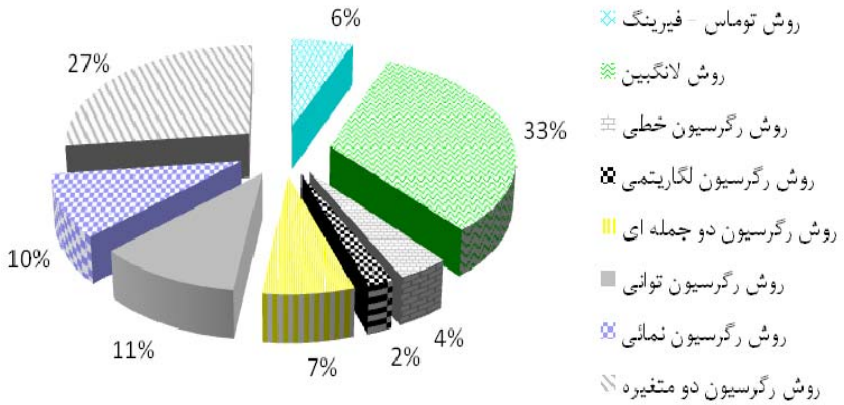
نتايج

براي اجراي روش‌هاي مختلف در هر ايستگاه از كل آمار ٢٠ درصد داده كنار گذاشته و فرض شد كه اين داده‌ها موجود نيست(به عبارت ديگر٨٠% داده‌ها براي كاليبراسيون و ٢٠% مابقي براي اعتبارسنجي)، سپس با استفاده از چهار روش در هر ايستگاه براي پارامترهاي دبي متوسط ماهانه، دبي متوسط سالانه، رسوب و كيفيت آب (شامل هدايت الكتريكي، كربنات، كلر، سولفات، مجموع آنيون، كلسيم، منيزيم، پتاسيم، مجموع كاتيون، درصد سديم، نسبت جذب سديم)، در ١٧ ايستگاه عمليات بازسازي داده‌ها انجام گرفت. در پايان، داده‌هاي برآوردي با داده‌هاي مشاهداتي كه كنار گذاشته شده، از طريق آماره اشتباه استاندارد با هم مقايسه شدند. از تجزيه و تحليل اطلاعات و نتايج به‌دست آمده از اعمال روش‌هاي مختلف بازسازي، موارد زير به عنوان نتايج اين پژوهش ارائه مي‌گردد.

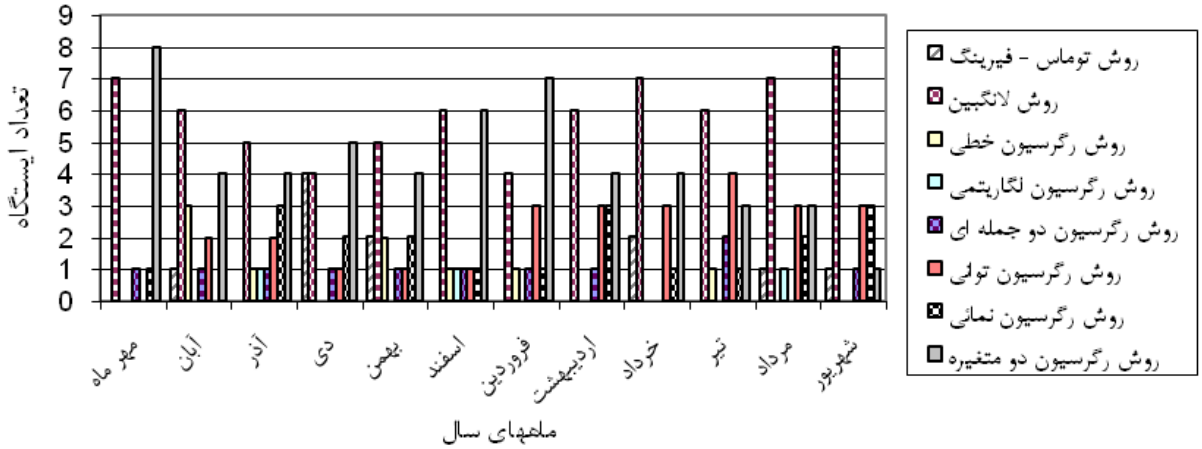
دبي متوسط ماهانه

در آزمون روش‌هاي مختلف بازسازي دبي متوسط ماهانه در ١٧ ايستگاه مطالعاتي، روش لانگبين بهترين وضعيت را براي بازسازي داده‌هاي دبي متوسط ماهانه نشان داد. در اين بررسي، از ١٩٢ مورد بازسازي دبي متوسط ماهانه، روش لانگبين با ٦٦ مورد(٣٣%) و روش رگرسيون دو متغيره با ٥٢ مورد(٢٧%) بهترين روش‌هاي بازسازي دبي متوسط ماهانه مي‌باشند. در مورد ساير روش‌ها، به دليل اينكه داراي اشتباه استاندارد بيشتري بوده و مقادير برآوردي با مقادير مشاهداتي تفاوت زيادي دارند، براي بازسازي دبي متوسط ماهانه در حوزه‌آبخيز سفيد رود مناسب نيستند. اين نتايج درشكل ٢ نشان داده شده است.

همچنين، در شكل ٣ روش‌هاي مختلف بازسازي دبي متوسط ماهانه براي ماه‌هاي مختلف سال به تفكيک نشان داده شده‌اند. از بررسي اشتباه استاندارد ماه‌هاي مختلف مشخص شد كه روش لانگبين در ماه‌هاي خرداد، تير، مرداد و شهريور كمترين مقدار اشتباه استاندارد را داشته است بنابراين، اين روش براي برآورد دبي در ماه‌هاي خشک در حوزه سفيد رود مناسب‌تر است.



شكل ۲- نمودار درصد كارايي روش‌هاي مختلف بازسازي دبي متوسط ماهانه



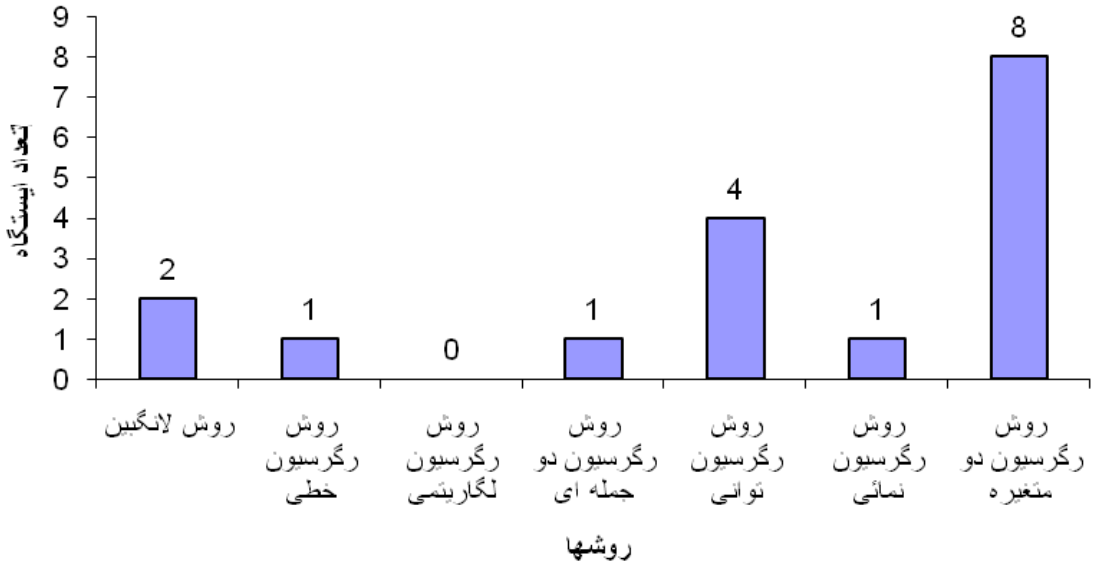
شكل ۳- نمودار مقايسه روش‌هاي بازسازي داده‌هاي دبي متوسط ماهانه در ايستگاه‌هاي مطالعاتي

دبي متوسط سالانه

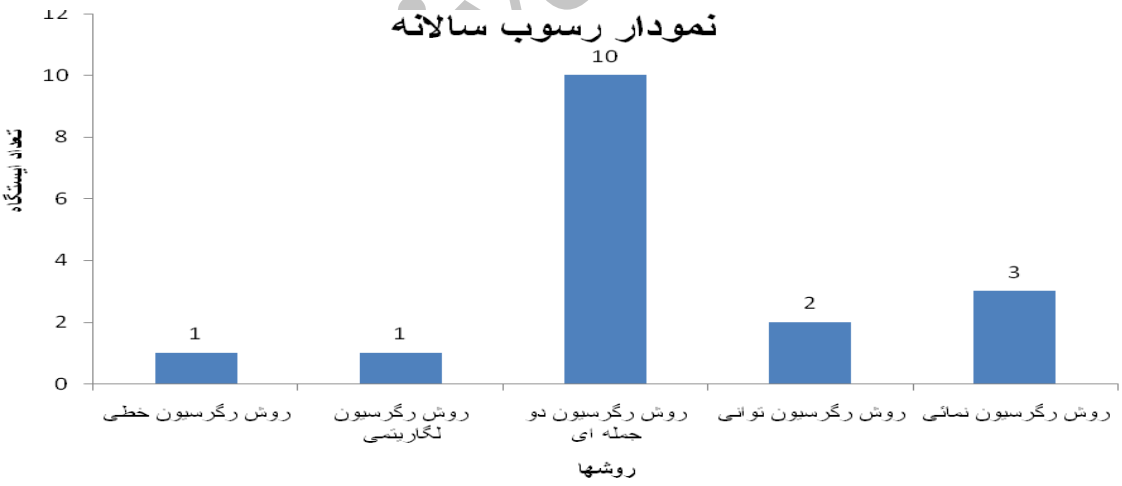
نتايج بررسي مقادير اشتباه استاندارد محاسبه شده براي دبي متوسط سالانه برآوردي نشان مي‌دهد كه روش رگرسيون دو متغيره از بين ۱۷ ايستگاه در ۸ ايستگاه بهترين جواب را داده است، بنابراين براي بازسازي دبي متوسط سالانه در حوزه سفيدرود روش مناسبي است (شكل ۴).

رسوب

براي بازسازي داده‌هاي رسوب در ايستگاه‌هاي مطالعاتي تنها از روش‌هاي رگرسيون تک متغيره استفاده شده است، چون كه ساير روش‌ها داراي اشتباه استاندارد بالائي بودند و بنابراين بهتر است از هيچ كدام از اين سه روش در بازسازي داده‌هاي رسوب ايستگاه‌هاي حوزه آبخيز سفيدرود استفاده نشود و ليكن از بين روش‌هاي مختلف رگرسيون تک متغيره، روش رگرسيون دوجمله‌اي بهتر از ساير روش‌ها بوده و در بين روش‌هاي مختلف رگرسيون تک متغيره مقدار اشتباه استاندارد كمتري دارد و اختلاف اكثر برآوردهاي رسوب با اين روش از مقادير مشاهداتي در محدوده ۱۰ درصد قرار مي‌گيرد. لازم به ذكر است كه اين داده‌ها از نظر پرت بودن كنترل شدند كه مشخص شد در سطح اطمينان ۱۰ درصد جزء داده پرت به حساب نمي‌آيند. پس در كل مي‌توان نتيجه گرفت كه در بازسازي داده‌هاي رسوب در ايستگاه‌هاي حوزه آبخيز سفيدرود بهتر است از روش‌هاي ديگري استفاده شود (شكل ۵).



شكل ۴- نمودار مقايسه روش‌هاي بازسازي داده‌هاي دبي متوسط سالانه در ايستگاه‌هاي مطالعاتي

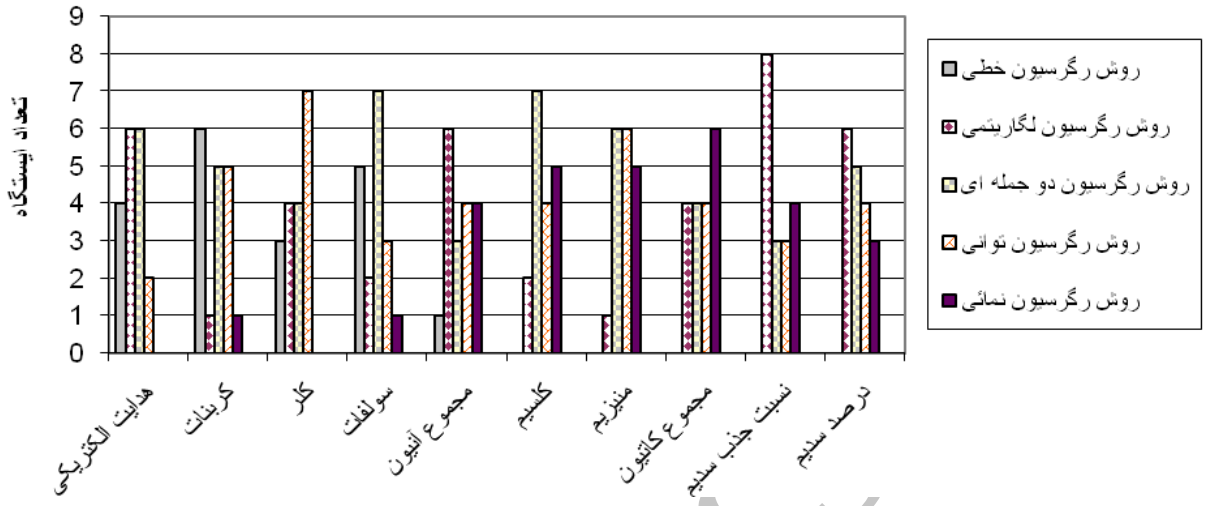


شكل ۵- نمودار مقايسه روش‌هاي بازسازي داده‌هاي رسوب متوسط سالانه در ايستگاه‌هاي مطالعاتي

كيفيت آب

چون ساير روش‌ها داراي اشتباه استاندارد بالائي بودند. براي بازسازي داده‌هاي كيفيت آب نيز مشابه بازسازي داده‌هاي رسوب عمل شد (تنها روش‌هاي رگرسيون تک متغيره). هر چند در مورد برخي پارامترها مانند مقدار كلسيم و منيزيم روش رگرسيون دو جمله‌اي برآوردهاي مناسبي را ارائه داده است و مقدار اشتباه استاندارد آن نيز پائين مي‌باشد، اما به‌طوركلي مي‌توان گفت كه از روش‌هاي رگرسيون نيز نمي توان براي بازسازي داده‌هاي كيفيت آب استفاده كرد. دليل آن تاثير پارامترهاي ديگر مانند شكل فرسايش غالب در منطقه، وجود تشكيلات آهكي، گچي و مارني در منطقه، چشمه‌هاي معدني، نوع

بارش، آلودگي‌ها و... روي كيفيت آب مي‌باشد. پس بهتر است كه براي بازسازي داده‌هاي كيفيت آب نيز روش‌هاي ديگري انجام و در صورت دستيابي به نتيجه مناسب استفاده شوند (شكل ۶).



شكل ۶- نمودار مقايسه روش‌هاي بازسازي داده‌هاي كيفيت آب در ايستگاه‌ها

بحث و نتيجه‌گيري

با توجه به اينكه به منظور برنامه‌ريزي و مديريت منابع آب، آبخيزداري، پيش‌بيني سيل، طراحي و عمليات ذخيره، آبرساني، كشتيراني، تفريح و مديريت زيست محيطي آگاهي از هيدرومتري ضروري است. كامل بودن اين آمار از اهميت بالايي برخوردار مي‌باشد. متاسفانه آمارهاي موجود در كشور ما عمدتاً داراي نواقص آماري مي‌باشند كه در مطالعات موجب ايجاد اشكالاتي مي‌شود و به ناچار بايد اقدام به رفع اين نواقص كرد. در اين بررسي سعي شد روش‌هاي مناسب براي بازسازي داده‌هاي هيدرومتري ارئه شود.. نتايج حاصل از اين پژوهش با نتايجي كه كه (Domokos, 1984 & Kovac) و (Clarke, 1979) به‌دست آورند تا حدودي مطابقت دارد زيرا ايشان روش رگرسيون دو متغيره را مناسب تشخيص داده بودند كه با نتايج اين پژوهش مطابقت دارد. نتيجه اين پژوهش با ساير منابع مطالعه شده مطابقت ندارد. به عنوان مثال (Sen, 1978) به اين نتيجه رسيد كه روش توماس فيرينگ براي گسترش داده‌هاي دبي ماهانه مناسب است كه با نتيجه اين پژوهش مطابقت ندارد كه اين مورد به دليل متفاوت بودن اقليم منطقه مورد مطالعه توسط نامبردگان با حوزه سفيدرود مي‌تواند باشد. چون روش توماس فيرينگ براي دبي‌هاي صفر مقادير را منفي برآورد مي‌كند كه مقدار دبي نمي‌تواند منفي باشد.

براي بازسازي داده‌هاي دبي متوسط ماهانه در حوزه‌آبخيز سفيدرود پيشنهاد مي‌شود كه از دو روش رگرسيون دو متغيره و روش لانگبين به صورت هم‌زمان استفاده شود. بنابراين براي ماه‌هايي از سال كه مقادير دبي صفر مي‌باشد بهتر است از روش لانگبين استفاده نشود و از روش رگرسيون دوجمله‌اي استفاده شود. همچنين ساير روش‌ها (مانند رگرسيون لگاريتمي) چون دبي متوسط ماهانه را خيلي بيشتر از واقعيت برآورد مي‌كنند و در بعضي حالات حتي مقادير دبي را منفي برآورد مي‌نمايند بهتر است از اين روش‌ها در بازسازي دبي‌هاي متوسط ماهانه استفاده نشود.

نتايج اين پژوهش نشان داد كه روش توماس فيرينگ نسبت به ساير روش‌ها از دقت مناسبي برخوردار نبوده و همچنين بين روش‌هاي مختلف بازسازي داده‌هاي هيدرومتري تفاوت معني‌داري وجود دارد. هر چند اين تفاوت بين دو روش لانگبين و رگرسيون دو متغيره كمتر

است. در پايان لازم به ذكر است كه فرضياتي كه در ابتداي پژوهش در نظر گرفته شده بود، رد مي‌شوند.

References

- Alizadeh, A. 2003. application hydrology, ver.15, Astan- e- ghodse Razavi Press.812pp.

-Ashraf Zadeh, A., Kholghi, M. and Mousavi Nadooshan, S. 2005. Discharg has produced long-term monthly average (case study: Dez River Basin) Iranian Journal of Natural Resources 58 (3), 493 - 501.

- Cadavid, L.G., Salas, J.D. and Boes, D.C. 1992. Disaggregation of short-term precipitation series. in Water Resources Papers. Vol 106. Colorado State University. Fort Collins CO.

- Clarke, R.T. 1979. Extension of Annual stream flow Record by correlation with precipitation subject to Hetergeneous Errors. Water Resources Research 15(5),1081-1088.

- Domokos, M. and Kovac, S.G. 1984. Extension of stream flow time series by Regression. Deutsche Gewasser kundliche mittellungen 28 (3), 85-89.

- Hisdal, H., Tveito, O.E. 1993. Extension of runoff series using empirical orthogonal functions. Hydrological Sciences journal 38(1), 33-49.

- Mahdavi, M. 2004. application hydrology, Vol 2, University of Tehran. Press. 312p.

- Naghdi, R., Sadatinejad, S. J. and Shayannejad, M. 2006.Comparison of two methods of multivariate linear regression and autocorrelation in the reconstruction monthly discharge data (case study : Ahwaz hydrometric station). Regional Conference First utilization of water resources Zayandehrud and Karun basins. 5pp.

- Patra, K.C. 2001. Hydrology and water Resources Engineering. Alpha science International LTD. 562 pp.

- SharifiNejhad, A. 1999. Estimated maximum moment Discharg with daily flow of some watershed jheophisic parameters (case study: Sefidrood watershed). Department of Natural Resources of Tehran University.

- Sharma, A., Lall, U. and Tarboton, D.G. 1997. Stream Flow simulation: A non parametric approach. Water Resources Research 33 (2), 291-308.

- Sheikh,V. B. 1999. Comparison of different reconstruction methods Discharg a maximum moment using the daily flow in the river Iran.. Department of Natural Resources, Tehran University.

- Sen, Z. 1978. A mathematical model of monthly flow sequences. Hydrological Sciences-Bulletin-des Sciences Hydrologiques 23 (6), 223-229.

Applicability of Various Reconstruction Methods of Hydrometric Data (Case Study: Sefidroud Basin)

H. Jalilian[1], Gh. Rostamizad[*2]and S. Arekhi[3]
[1] MSc Graduate, Faculty of Natural Resources, University of Tehran, Karaj, I.R. Iran
[2] Ph. D. Student Faculty of Natural Resources, University of Tehran, Karaj, I.R. Iran
[3] Assistant Professor, Faculty of Agriculture, Ilam University, Ilam, I.R. Iran

(Received:2010/December/28 , Accepted:2011/October/28)

Abstract

The necessity of using discharge data in hydrological designs and regional programming is an unavoidable matter. Unfortunately, malfunctioning of observation instruments, natural disasters and other problems sometimes result in incomplete or missing data. Having of data related to time series is necessary for every investigation about hydrometeorology. Thus missing data should be estimated in a proper way. There are different methods for generation of missing data which estimates the data with regard to particular parameters. In this study, we have used four methods including linear regression, multiple regression, Longbein method and Thomas-Fiering method for reconstruction of hydrometric data including monthly and annual discharge, sediment and water quality data. Results showed that among various methods, in 192 cases of reconstruction of monthly average discharge of 17 stations, Longbein and multiple regression methods in 33% and 27% of the cases have provided the best results, respectively. Two-variable regression in 8 of 17 stations had the best answer and it is suitable for estimating of annual average discharge. Also results indicated that we cannot use of relation between discharge-water quality and discharge-sediment to reconstruct data using above mentioned methods in the Sefidroud basin.

Keywords: Hydrometric data, Longbein method, Multiple regression, Sefidroud basin

*Corresponding author: Tel: +98 918 9409568 Fax: +982632249313 E-mail: gh_rostamizad@ut.ac.ir