تحقیقات منابع آب ایران
Iran-Water Resources Research

سال دهم، شماره ۳، زمستان ۱۳۹۳
Volume 10, No. 3, Winter 2015 (IR-WRR)
۹۹-۱۰۳

یادداشت فنی

تعیین مناطق همگن سیلابی به روش تحلیل خوشه‌ای افرازی K-means براساس گشتاورهای خطی مرتبه دوم (مطالعه موردی: حوضه سدهای رودبار و بختیاری)

رسول شاهسون[۱*]، علی آهنی[۲]، سید سعید موسوی ندوشنی[۳] و علی مریدی[۴]

چکیده

گروهی از سایت‌ها با همگنی کافی در فرآیندهای تولید سیلاب، یک منطقه‌ی همگن برای تحلیل فراوانی منطقه‌ای سیلاب را تشکیل می‌دهند. تحلیل خوشه‌ای عنوان گونه‌ای از روش‌های تحلیل چند متغیری آماری برای گروه‌بندی یک سری داده به چندین گروه یا خوشه‌ها است. این روش، روش کارآمدی برای تحلیل فراوانی منطقه‌ای سیلاب است. در این گزارش با استفاده از زبان برنامه‌نویسیR، تحلیل خوشه‌ای بر اساس گشتاورهای خطی مرتبه دوم داده‌های حداکثر لحظه‌ای ۳۴ ایستگاه هیدرومتری واقع در استان لرستان و در محدوده حوضه سدهای رودبار و بختیاری برای تعداد ۲ تا ۶ خوشه صورت گرفته است. در تحلیل خوشه‌ای، از الگوریتم K-means که در نوع خوشه‌بندی افرازی دسته‌بندی می‌شود، بهره گرفته شده است. گروه‌های ادغام شده از ایستگاه‌ها تشکیل شده و سپس با انجام آزمون همگنی گروه‌های همگن حوضه سدهای رودبار و بختیاری تعیین شده‌اند.

کلمات کلیدی: منطقه‌بندی، تحلیل خوشه‌ای، گشتاورهای خطی، K-means، سدهای رودبار و بختیاری.

تاریخ دریافت مقاله: ۱۳ اردیبهشت ۱۳۹۲
تاریخ پذیرش مقاله: ۱۳ آبان ۱۳۹۳

۱- دانش‌آموخته‌ی کارشناسی ارشد مهندسی رودخانه، پردیس فنی و مهندسی شهید عباسپور، دانشگاه شهید بهشتی، تهران
۲-دانش‌آموخته‌ی کارشناسی ارشد سازه‌های آبی، دانشگاه صنعتی شاهرود
۳- استادیار پردیس فنی و مهندسی شهید عباسپور، دانشگاه شهید بهشتی، تهران
۴- استادیار پردیس فنی و مهندسی شهید عباسپور، دانشگاه شهید بهشتی، تهران
*- نویسنده مسئول

Technical Note

Recognition of Homogeneous Regions Using K-means Partitional Cluster Analysis Method Based on Second-order L-moment (Case Study: Roodbar Dam and Bakhtiari Dam Basins)

R. Shahsavan [1] , A. Ahani [2], SS. Mousavi Nadoushani [3] and A. Moridi [4]*

Abstract

A group of sites with sufficient homogeneity in producing flood mechanisms form a homogeneous region for regional flood frequency analysis. Cluster analysis is the generic name of a variety of multivariate statistical procedures that are used to classify given data into similar groups or clusters. This procedure is an efficient method for regional flood frequency analysis. In this article, cluster analysis of annual maximum flow data of 34 gauging stations located in Lorestan province in Roodbar Dam and Bakhtiari Dam basins for 2 to 6 clusters has been performed using R programming language. K-means algorithm that is a type of partitional clustering has been used for cluster analysis. Merged groups of sites have been formed and homogeneous groups of Roodbar and Bakhtiari dam basins has then been identified by performing homogeneity test.

Keywords: Regionalization, Cluster Analysis, L-moments, K-means, Roodbar and Bakhtiari Dams.

Received: May 3, 2014
Accepted: October 5, 2014

1- M.Sc. of River Engineering, Shahid Abbaspour Campus, ShahidBeheshti University, Tehran, Iran. Email: shahsavanrasoul@yahoo.com
2 M.Sc. of Agricultural Engineering – Water Structures, School of Agricultural Engineering, Shahrood University of Technology, Shahrood, Iran
2 Assistant Professor, Faculty of Water and Environmental Engineering, Shahid Abbaspour Campus, Shahid Beheshti University, Tehran, Iran.
2 Assistant Professor, Faculty of Water and Environmental Engineering, Shahid Abbaspour Campus, Shahid Beheshti University, Tehran, Iran.
*- Corresponding Author

۱- مقدمه

روش‌های متعددی برای منطقه‌بندی حوضه‌های آبریز و یافتن مناطق همگن هیدرولوژیک توسعه یافته‌اند که شباهت بین سایت‌ها را از طریق آزمودن ویژگی‌های حوضه مانند مشخصات فیزیوگرافیک، موقعیت جغرافیایی، ویژگی‌های هواشناسی، مشخصات زمین‌شناسی، آماره‌های درون‌سایتی سیلاب و ... جست و جو می‌کنند. هر روش منطقه‌بندی برتری‌ها و محدودیت‌های خاص خود را دارد. با این حال، به دلیل محدودیت‌های ناشی از کمبود داده‌ها و تأثیر قضاوت‌های ذهنی در انتخاب مشخصات، وزن‌ها، مقادیر آستانه و شاخص‌های فاصله، هیچ معیار قطعی که توسط آن برتری هر یک از روش‌های منطقه‌بندی بتواند به روشنی بیان شود، وجود ندارد (Rao and Srinivas, 2008).

تحلیل خوشه‌ای[1] نام گونه‌ای از روش‌های آماری چند متغیری است که به منظور تحقیق، تفسیر، و دسته‌بندی داده‌های موجود در گروه‌های مشابه یا خوشه‌هایی که می‌توانند هم‌پوشانی داشته باشند یا نداشته باشند، مورد استفاده قرار می‌گیرند. نقاط معرف داده‌ها در یک خوشه باید تا حد امکان مشابه باشند و نقاط معرف داده‌ها در خوشه‌های مختلف باید حتی‌الامکان متفاوت باشند. روش‌های منطقه‌بندی مبتنی بر تحلیل خوشه‌ای می‌توانند در تشخیص گروه‌های حوضه‌هایی که فرآیند تولید سیلاب مشابهی دارند، نقش مؤثری را ایفا کنند. پژوهش‌ها نشان می‌دهند قضاوت‌های ذهنی و تلاش مورد نیاز برای تعیین گروه‌های حوضه‌های آبریز با استفاده از روش‌های معمول، با به کارگیری شیوه‌های خوشه‌بندی کارآمد به شکل چشم‌گیری کاهش می‌یابد (Rao and Srinivas, 2008).

در این مقاله عملکرد الگوریتم خوشه‌بندی K-means در تشکیل مناطق همگن سیلابی در حوضه‌های سدهای رودبار و بختیاری مورد بررسی قرار گرفته است. همچنین در این مطالعه گشتاورهای خطی مرتبه‌ی دوم داده‌های دبی حداکثر لحظه‌ای ثبت شده در ایستگاه‌های هیدرومتری به‌عنوان ویژگی‌ها یا بردارهای مشخصه‌ی نقاط داده‌ها در خوشه‌بندی مورد استفاده قرار گرفته‌اند تا بدین‌ترتیب کارایی این ویژگی در تشکیل مناطق همگن سیلابی آزموده شود.

۲- روش تحقیق

در این گزارش از الگوریتم خوشه‌بندی K-means در تحلیل خوشه‌ای بهره گرفته شده است. این الگوریتم در روش خوشه‌بندی افرازی دسته‌بندی می‌شود. در روش K-means، هر خوشه توسط مرکز خود که میانگین وزنی یا غیر وزنی بردارهای مشخصه‌ی موجود در خوشه است، معرفی می‌شود. این روش به دلیل کارایی خود در خوشه‌بندی مجموعه‌های بزرگ داده‌ها با ویژگی‌های عددی مشهور است. به علاوه، این روش نسبت به وجود داده‌های پرت حساس است. انتخاب اولیه‌ی مراکز خوشه‌ها به صورت بهینه، یکی از مسائل مهم در استفاده از این الگوریتم است (آهنی و همکاران، ۱۳۹۲).

با کمینه‌سازی تابع هدف، فاصله‌ی هر بردار مشخصه از مرکز خوشه‌ای که به آن تعلق دارد به حداقل می‌رسد. در الگوریتم عددی K-means، بردارهای مشخصه در تکرارهای متعدد از خوشه‌ای به خوشه‌ی دیگر انتقال می‌یابند تا مقدار تابع هدف کمینه شود (Rao and Srinivas, 2008).

برای ارزیابی همگنی مناطق تشکیل شده از شاخص ناهمگنی H_1 مبتنی بر گشتاورهای خطی استفاده می‌شود (Hosking and Wallis, 1997). مطابق جدول ۱، تمام مناطق دارای مقدار H_1 کمتر از ۲ برای تعیین مناطق همگن مورد استفاده قرار می‌گیرد.

جدول ۱- تفسیر معیار همگنی

معیار همگنی H_1	تفسیر	بازبینی
< ۱	همگن	نیازی نیست
۱-۲	احتمالاً همگن	اختیاری
۲-۴	ناهمگن	مطلوب
> ۴	قویاً ناهمگن	ضروری است

طبق اظهار Hosking and Wallis (1997) انواع تعدیلات مانند جابه‌جایی یک یا تعداد اندکی سایت از منطقه‌ای به منطقه‌ی دیگر، حذف یک یا تعداد کمی از سایت‌ها از سری داده‌ها، تقسیم کردن منطقه، ادغام یک منطقه با مناطق دیگر،به‌دست آوردن داده‌های بیشتر و بازتعریف گرهٔ است برای بهبود ارتباط فیزیکی مناطق و کاهش ناهمگنی مناطق مفید باشند.

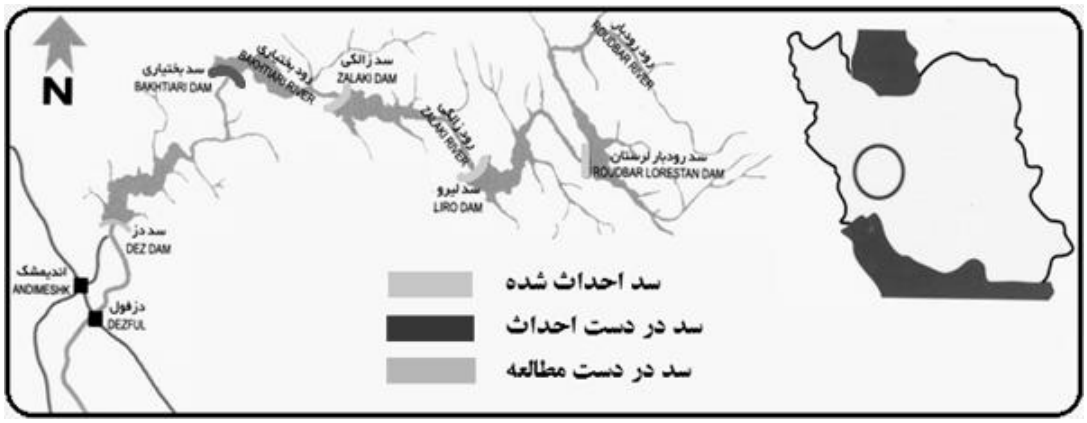
حوضه آبریز رودخانه بختیاری در جنوب غرب کشور در محدوده مختصات جغرافیایی '۱۵ °۴۸ تا '۲۰ °۵۰ طول شرقی و '۳۰ °۳۲ الی '۳۰ °۳۳ عرض شمالی قرار گرفته است. مهم‌ترین ایستگاه آبسنجی رودخانه تنگ پنج نام دارد که نزدیک محل الحاق رودخانه سزار واقع شده است. محل پیشنهادی سد مخزنی بختیاری حدود ۲/۸ کیلومتر در بالادست ایستگاه آبسنجی مزبور خواهد بود (شرکت توسعه منابع آب و نیروی ایران، ۱۳۸۵a).

سد مخزنی رودبار واقع در جنوب شهر الیگودرز در استان لرستان و بر روی رودخانه رودبار از سرشاخه‌های دز و با موقعیت جغرافیایی ′۴۹°۴۱ طول شرقی و ′۳۲°۵۴ عرض شمالی و رقوم محور ۱۶۲۰ متر از سطح دریا هم اکنون در فاز مطالعات تکمیلی و مراحل اولیه اجرایی قرار دارد. مساحت حوضه آبریز رودخانه تا محل سد ۲۲۵۵ کیلومترمربع می‌باشد. با توجه به توزیع ارتفاعی محدوده مورد مطالعه حدود ۲۲ درصد مناطق، رقومی بالاتر از ۲۵۰۰ متر از سطح دریا دارند و شیب متوسط حوضه تا محل سد ۲۷/۷ درصد است (شرکت توسعه منابع آب و نیروی ایران، ۱۳۸۵b). شکل ۱ موقعیت سدهای رودبار و بختیاری نسبت به سدهای مجاور و شکل ۲ موقعیت ۳۴ ایستگاه هیدرومتری مورد مطالعه واقع در استان لرستان را نشان می‌دهد.

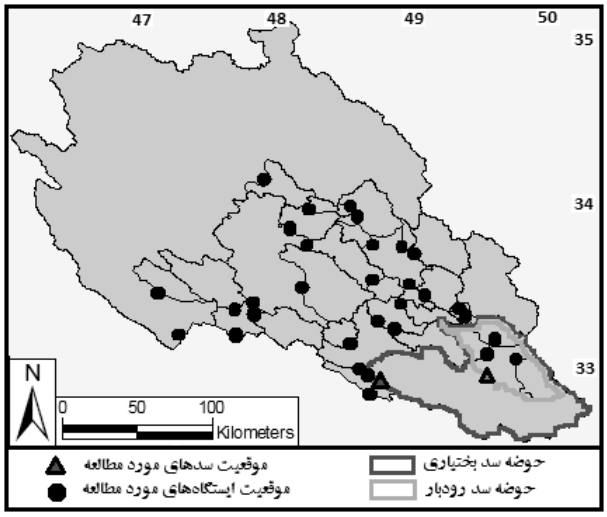
داده‌های مورد استفاده در این گزارش از بخش آمار ایستگاه‌های هیدرومتری شرکت آب منطقه‌ای استان لرستان تهیه شده است.

۳- نتایج و بحث

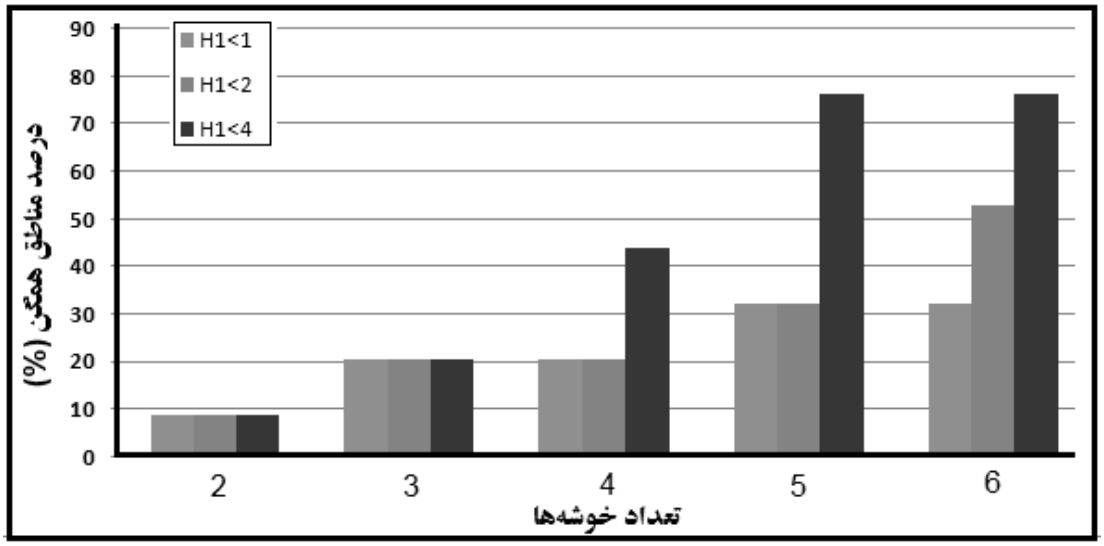
با استفاده از تحلیل خوشه‌ای بر اساس گشتاورهای خطی آمار حداکثر لحظه‌ای سالانه ایستگاه‌های هیدرومتری در دوره مشترک آماری ۳۰ ساله، ایستگاه‌های ادغام شونده در هر خوشه تعیین گردید. تحلیل‌ها برای ۲ تا ۶ خوشه صورت گرفت که نتایج حاصل در جدول ۲ آورده شده است. با انجام آزمون همگنی مقدار H_1 هر خوشه مشخص شد از میان خوشه‌های مختلف در تحلیل‌های ۲ تا ۶ خوشه‌ای، خوشه‌های با معیار همگنی $H_1<2$ همگن تلقی می‌شوند. در جدول ۳ خوشه‌های همگن سدهای مورد مطالعه رودبار و بختیاری آورده شده‌اند. پس از تعیین ایستگاه‌های خوشه‌های مختلف در تحلیل خوشه‌ای به کمک گشتاورهای خطی و انجام آزمون همگنی، درصد مناطق همگن در حالت‌های ۲ تا ۶ خوشه‌ای محاسبه گردید. نتایج این محاسبات برای سه مقدار مختلف $H_1<1$، $H_1<2$ و $H_1<4$ در شکل ۳ نشان داده شده است.



شکل ۱- موقعیت سدهای رودبار و بختیاری نسبت به سدهای مجاور



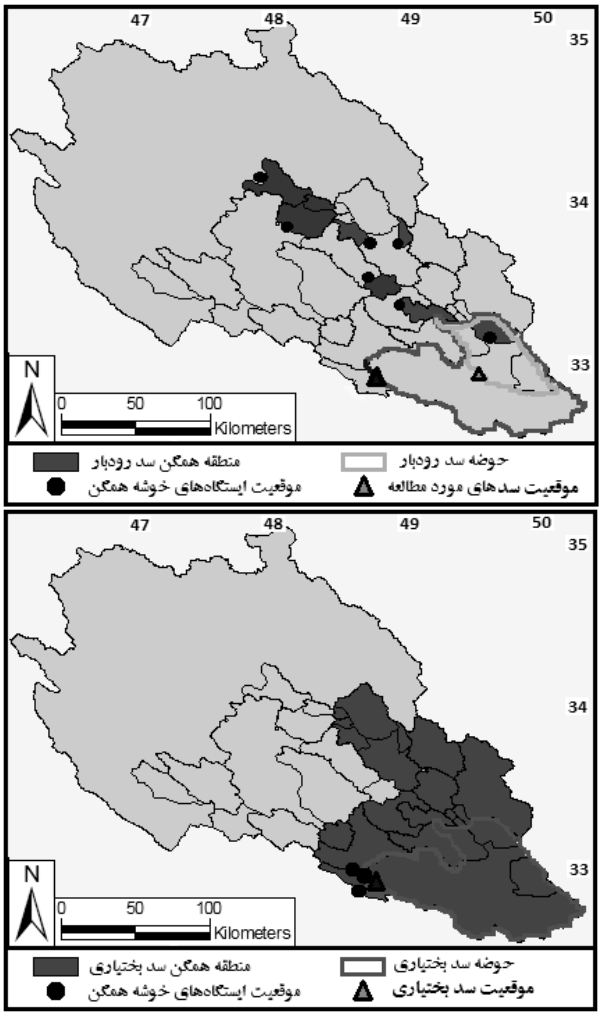
شکل ۲- موقعیت ۳۴ ایستگاه مورد مطالعه و حوضه بالادست آن‌ها



شکل۳- درصد مناطق همگن تحلیل خوشه‌ای به کمک گشتاورهای خطی در مقابل تعداد خوشه‌های مختلف

با انجام تحلیل خوشه‌ای بر اساس گشتاورهای خطی مرتبه دوم داده‌های حداکثر سالانه لحظه‌ای برای ۳۴ ایستگاه مورد مطالعه، منطقه همگن سدهای مورد مطالعه تعیین گردید. از آن‌جا که ایستگاه تنگ پنج بختیاری بلافاصله پس از موقعیت سد بختیاری قرار دارد، در خوشه‌های حاصل از تحلیل خوشه‌ای، ایستگاه مذکور به عنوان ایستگاه نماینده سد بختیاری در نظر گرفته شد. از آن‌جا که در حوضه سد رودبار ایستگاهی نزدیک موقعیت سد وجود ندارد، برای سد رودبار، خوشه‌های همگنی که شامل هر یک از سه ایستگاه هیدرومتری بالادست موقعیت سد یعنی زردفهره، سکانه و کاظم آباد هستند در تعیین بهترین منطقه همگن لحاظ شده است.

برای سدهای مورد مطالعه بختیاری و رودبار لرستان، نتایج حاصل از بهترین خوشه‌های همگن از لحاظ تعداد ایستگاه‌های ادغام شده در یک گروه همگن در شکل ۴ نشان داده شده است.



شکل۴- مناطق همگن سدهای رودبار و بختیاری

۴- خلاصه و جمع‌بندی

در این مقاله با انجام تحلیل خوشه‌ای بر اساس گشتاورهای خطی مرتبه دوم داده‌های حداکثر سالانه لحظه‌ای برای ۳۴ ایستگاه مورد مطالعه واقع در استان لرستان، منطقه همگن سدهای رودبار و بختیاری تعیین گردید. منطقه همگن سد بختیاری شامل سه ایستگاه هیدرومتری تنگ پنج سزار، تنگ پنج بختیاری و تله زنگ می‌باشد، که در ۵ خوشه مختلف همگن در حالات ۲، ۳، ۴، ۵ و ۶ خوشه‌ای حضور دارند. همچنین نتایج حاصل از تحلیل خوشه‌ای، حاکی از آن است که بهترین منطقه همگن سد رودبار لرستان شامل ۷ ایستگـــاه

هیدرومتری نورآباد، دهنو، سراب سید علی، برآفتاب، چم چیت، کاظم آباد و تنگ محمد حجی در حالت ۶ خوشه‌ای می‌باشد. از داده‌های مناطق همگن سدهای مورد مطالعه حاصل تحلیل خوشه‌ای، می‌توان در تحلیل منطقه‌ای سیلاب بهره گرفت. نتایج بررسی تحلیل خوشه‌ای برای حالات ۲ تا ۶ خوشه‌ای نشان می‌دهد که با افزایش تعداد خوشه‌ها از ۲ تا ۶ درصد مناطق همگن نیز افزایش می‌یابد. همچنین با لحاظ مقدار آستانه H_1 از ۱ تا ۴ تعداد خوشه‌های همگن و در نتیجه درصد مناطق همگن افزایش خواهد یافت.

جدول ۲- نتایج حاصل از آزمون همگنی برای حالت‌های مختلف خوشه‌بندی

تعداد خوشه‌ها	شماره‌ی خوشه	تعداد ایستگاه‌ها	H_1
۲	۱	۳۱	۶/۲۵
	۲	۳	-۰/۸۱
۳	۱	۲۷	۶/۳۸
	۲	۳	-۰/۸۱
	۳	۴	-۰/۱۲
۴	۱	۳	-۰/۸۱
	۲	۴	-۰/۱۶
	۳	۸	۳/۱۹
	۴	۱۹	۵/۸۴
۵	۱	۸	۴/۶۹
	۲	۴	-۰/۵۱
	۳	۱۵	۳/۹۳
	۴	۳	-۰/۸۱
	۵	۴	-۰/۰۹
۶	۱	۳	-۰/۸۱
	۲	۷	۱/۱۳
	۳	۴	-۰/۰۵
	۴	۴	-۰/۵۱
	۵	۸	۴/۸۶
	۶	۸	۳/۰۱

جدول ۳- خوشه‌ی همگن سدهای رودبار و بختیاری

گروه ادغام شده از ایستگاه‌ها	تعداد ایستگاه	H_1	خوشه	سد
۲۹۳-، ۲۱-۲۹۵ ۲۱-۲۹۱ ، ۲۱	۳	-۰/۸۱	۱-۶ ، ۴-۵، ۱-۴، ۲-۳، ۲-۲	بختیاری
۴۰۰-، ۲۱-۴۶۱ ۲۱، ۲۱-۲۸۱ ، ۱۷۱- ۲۱-۱۸۱ ۲۱، ۲۱-۱۶۷، ۲۱-۱۴۵	۷	۱/۱۳	۲-۶	رودبار

پی نوشت

1-Cluster Analysis

۵- تشکر و قدردانی

نسخه خلاصه‌ای از این مقاله در پنجمین کنفرانس مدیریت منابع آب ایران ارائه شده است.

۶- مراجع

آهنی ع، امامقلی‌زاده ص، موسوی ندوشنی س، اژدری خ (۱۳۹۲) منطقه‌بندی حوضه‌ی آبریز سفیدرود بزرگ با استفاده از روش‌های خوشه‌بندی ترکیبی. مجموعه مقالات هفتمین کنگره ملی مهندسی عمران، زاهدان، ایران، ۱۸-۱۷ اردیبهشت.

شرکت توسعه منابع آب و نیروی ایران (۱۳۸۵a) گزارش تکمیلی هیدرولوژی طرح سد و نیروگاه برقابی بختیاری.

شرکت توسعه منابع آب و نیروی ایران (۱۳۸۵b) گزارش هیدرولوژی طرح سد و نیروگاه برقابی رودبار.

Hosking JRM, Wallis JR (1997) Regional frequency analysis: An approach based on l-moments, Cambridge University Press, London, 224p.

Rao AR, Srinivas VV (2008) Regionalization of watersheds: An approach based on cluster analysis, Springer Publishers, 248p.