

بررسی و آنالیز مدل آسیب پذیری آبخوان با استفاده از توابع متعامد تجربی و تحلیل خوشه‌ای

عمار حبیبی کندبن^{۱*}، رزگار عربزاده^۲، افشین اشرفزاده^۳

۱. دانشجوی کارشناسی مهندسی آب دانشگاه گیلان

۲. دانشجوی کارشناسی ارشد مهندسی منابع آب دانشگاه تهران

۳. استادیار گروه مهندسی آب دانشگاه گیلان

(تاریخ دریافت: ۱۳۹۴/۹/۲۹ - تاریخ تصویب: ۱۳۹۵/۱/۲۸)

چکیده

یکی از مدل‌های متداول در برآورد میزان آسیب‌پذیری آبخوان، مدل DRASTIC است. این مدل شامل هفت پارامتر، عمق تا سطح ایستابی، تغذیه خالص به آبخوان، زمین‌شناسی ناحیه اشباع، قشر خاک سطحی، شیب، اثر ناحیه غیراشباع و هدایت هیدرولیکی آبخوان است. در این مقاله پارامتر این مدل استخراج گردیده و درازای استفاده از ترکیب خطی لایه‌ها جهت تقریب میزان آسیب‌پذیری از تحلیل مؤلفه‌ها اساسی (PCA) به‌عنوان روش جایگزین استفاده شده است. مزیت این روش برآورد اوزان مناسب و اصلاح‌شده لایه‌ها و تخمین مناسب میزان آسیب‌پذیری با توجه به در نظر گرفتن واریانس حداکثری مؤلفه‌های مدل آسیب‌پذیری است. در مطالعه صورت گرفته مدل آسیب‌پذیری آبخوان دشت قره-دهگلان با استفاده از روش تحلیل مؤلفه‌های اساسی استخراج و به‌وسیله متدهای تحلیل خوشه‌ای (CA) تفکیک گردید. درنهایت نیز با استفاده از روش ارزیابی، کارایی خوشه‌بندی Dunn، اقسام خوشه‌بندی‌های صورت گرفته مورد آزمون قرار گرفت و بهترین روش نیز معرفی گردید. در این مطالعه مؤلفه اول ۷۲/۵ درصد از واریانس کل داده‌ها را توجیه نموده و روش خوشه‌بندی Single نیز به‌عنوان بهترین روش برای کاربردی کردن نتایج حاصل از این تحقیق معرفی گردید.

واژه‌های کلیدی: آبخوان، آسیب‌پذیری، PCA، CA، قره-دهگلان

مقدمه

با توسعه جوامع بشری و گسترش فعالیت‌های صنعتی سهم عمده‌ای در آلودگی‌های محیط‌زیست به‌ویژه آب دارد (Rahman, 2008)؛ بنابراین فعالیت‌های انسان در بعضی از موارد به توازن‌های ظریف موجود در طبیعت آسیب می‌رساند بعضی از این آسیب‌ها توسط طبیعت قابل جبران هستند. به عبارتی طبیعت قادر است بعضی از خرابی‌ها را ترمیم نماید؛ اما توانایی جبران آلودگی در طبیعت شامل تمام آسیب‌های وارده به آن نمی‌شود و نیز جبران خسارت وارده به زمان زیاد نیاز دارد؛ بنابراین، انسان باید با شناخت، پیشگیری و یا مقابله با آلودگی‌ها، آن‌ها را در کنترل خود درآورد. یکی از مهم‌ترین منابع آبی آسیب‌پذیر در برابر آلودگی، آبخوان‌ها می‌باشند. این منابع به شکل‌های مختلف در معرض آلودگی قرار دارند که تشخیص و کنترل آلودگی در آن‌ها نسبت به آب‌های سطحی مشکل‌تر و پرهزینه‌تر است. همچنین، به دلیل استمرار آلودگی در این منابع، بهترین روش جلوگیری از آلودگی آن‌ها، شناسایی

منابع آلوده‌کننده، مناطق آسیب‌پذیر، تهیه نقشه‌های پهنه‌بندی آسیب‌پذیری و اتخاذ سیاست‌های مدیریتی مناسب است (Babiker et al, 2005). تاکنون تعاریف زیادی در مورد آسیب‌پذیری آبخوان و مفهوم آن در هیدروژئولوژی ارائه شده است که از جمله می‌توان به تعریف کمیته ملی علوم آمریکا در سال ۱۹۹۳ اشاره کرد. این کمیته، آسیب‌پذیری آب‌های زیرزمینی نسبت به آلودگی را، تمایل یا احتمال رسیدن آلاینده‌ها به یک مکان مشخص در سیستم آب زیرزمینی بعد از به وجود آمدن در برخی محل‌ها در بالای سطح آبخوان می‌داند (Almasari, 2008). مفهوم آسیب‌پذیری آب‌های زیرزمینی برای اولین بار در فرانسه و در اواخر دهه ۱۹۶۰ به‌کاررفته است. امکان نفوذ و انتشار آلاینده‌ها به درون سیستم آب زیرزمینی را آسیب‌پذیری می‌نامند. آسیب‌پذیری به‌عنوان یک استعداد ذاتی سیستم منابع آب زیرزمینی در نظر گرفته می‌شود؛ که وابسته به میزان حساسیت این سیستم به تأثیرات انسانی و یا طبیعی است. آسیب‌پذیری ذاتی، به مفهوم حساسیت آبخوان به عوامل طبیعی است درحالی‌که آسیب‌پذیری ویژه، آسیب‌پذیری ذاتی را به همراه با احتمال قرار گرفتن آب زیرزمینی در معرض نفوذ آلاینده‌ها موردبررسی قرار می‌دهد. ولی از نظر متخصصین میان

* نویسنده مسئول: Ammarhabibi@ut.ac.ir

زیادتری را مورد بررسی قرار داد. همچنین افراد زیادی از این روش‌ها جهت مطالعات منابع آب و سایر تحلیل‌های علمی خود استفاده کرده‌اند. Thareja *et al.* (2011) با مطالعه‌ای که با روش PCA بر روی رودخانه Ganga در شهرستان کانپور انجام دادند. ۱۴ متغیر کیفیت آب (فیزیکی و شیمیایی) را مورد بررسی قرار دادند به این نتیجه رسیدند که دو عامل اثر انسانی و فعالیت‌های صنعتی بیشترین تأثیر را بر کیفیت آب این منطقه گذاشته‌اند. Nasir *et al.* (2011) با انجام مطالعه روی رودخانه Klang واقع در مالزی با استفاده از مدل رگرسیون خطی و PCA، عواملی همچون فرسایش خاک، ورودی انسانی، رواناب سطحی، مواد شوینده، زباله‌های شهری، پساب صنعتی و زباله‌های مسکونی تأثیر بسیاری در کیفیت آب این منطقه ایفا می‌کنند که در بین این عوامل، آلودگی داخل شهری به‌عنوان بالاترین عامل مؤثر در آلودگی رودخانه Klang شناخته شد. Wang *et al.* (2013) در مطالعه‌ای که بر روی رودخانه Songhua واقع در هاربین چین با استفاده از CA^۲ و PCA انجام دادند، به این نتیجه رسیدند که شناسایی خطرات زیست‌محیطی در مناطق مختلف این مطالعه به مدیران مسئول کمک خواهد کرد که تصمیم‌گیری بهتر و مناسب‌تری در چگونگی بهبود کیفیت آب داشته باشند.

هدف اصلی از مطالعه حاضر تهیه مدل آسیب‌پذیری آبخوان برای محل‌های مشاهداتی داده‌ها در محل لوگ چاه‌های اکتشافی است. این مطالعه شامل داده قبل از پردازش، تجزیه و تحلیل همبستگی و استخراج نتایج با استفاده از تجزیه و تحلیل مؤلفه‌های اساسی (PCA) برای مجموعه‌ای از داده‌ها و تخمین شاخص آسیب‌پذیری از طریق کاهش اجزا است. همچنین برای کاربردی کردن نتایج تحقیق از روش‌های خوشه‌بندی نیز استفاده شده است.

مواد و روش‌ها

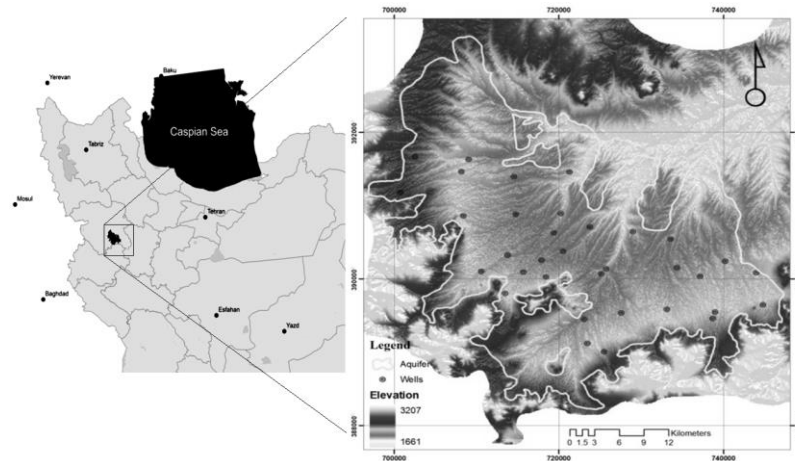
منطقه مورد مطالعه

دشت دهگلان بزرگ‌ترین دشت استان کردستان و مهم‌ترین منبع تولیدکننده محصولات کشاورزی است و متوسط سالانه بارش آن حدوداً ۳۵۰ میلی‌متر بوده و از نظر اقلیمی جزء نواحی سرد به شمار می‌آید (Ahmadi *et al.*, 2012). موقعیت ریاضی دشت شامل طول جغرافیایی و عرض جغرافیایی آن بین ۳۵ درجه و یک دقیقه تا ۳۵ درجه و ۳۹ دقیقه عرض شمالی نسبت به خط استوا و ۴۷ درجه و ۷ دقیقه تا ۴۷ درجه و ۳۶ دقیقه

این دو تعریف به‌صورت معین نمی‌توان تمایز روشنی قائل شد (National Academies Press, 1993). اهمیت بررسی آسیب‌پذیری از طرفی به خاطر این است که آب زیرزمینی یکی از منابع اصلی آب شرب است و از طرف دیگر تمرکز بالای بشر و فعالیت‌های اقتصادی به‌طور مثال صنعت، کشاورزی و مناطق مسکونی به‌عنوان منابع پتانسیل آلودگی آب زیرزمینی است و این موضوع در مناطق و شهرهای رو به گسترش از اهمیت بیشتری برخوردار است (Khodaei *et al.*, 2007). آبخوان‌ها به دلیل استعداد آلودگی کمتر و همچنین ظرفیت ذخیره زیاد نسبت به منابع آب‌های سطحی به‌عنوان یک منبع مهم در منابع آب مورد توجه است (Babiker *et al.*, 2005). وجود منابع مهم آلاینده‌های غیر نقطه‌ای و نقطه‌ای ناشی از فعالیت‌های انسانی در سطح زمین و نفوذ آن‌ها به آبخوان‌ها باعث کاهش کیفیت آب زیرزمینی می‌شود. به این خاطر بررسی و جلوگیری از آلودگی آب‌های زیرزمینی یکی از اقدامات مهم در مدیریت منابع آب زیرزمینی است (Fritch *et al.*, 2000). یکی از راه‌های مناسب برای جلوگیری از آلودگی آب زیرزمینی، تعیین میزان آسیب‌پذیری آبخوان و سوق دادن تلاش‌های مدیریتی به سمت این مناطق جهت حفظ کیفیت آب است (Khodaei, 2003). پس می‌توان گفت که آسیب‌پذیری به خصوصیات آبخوان و تنش‌های وارده به آن بستگی دارد و مشخصه‌ای نسبی، بدون بعد و غیرقابل اندازه‌گیری است. بررسی آب زیرزمینی نسبت به آلودگی و آسیب‌پذیری آن در طی سال‌های گذشته موضوع بسیاری از تحقیقات بوده است؛ و روش‌های مختلفی تاکنون جهت بررسی آن ابداع شده‌اند (Al - Secunda *et al.*, 1998- 1990). یکی از روش‌های مورد استفاده تحلیل مؤلفه‌های اساسی (PCA) است. در این مقاله برای بررسی آسیب‌پذیری آبخوان منطقه مورد مطالعه از روش PCA برای تجزیه و تحلیل استفاده شده است. مدت زیادی است که محققان از روش تحلیل مؤلفه‌های اساسی در علوم مختلف استفاده کرده‌اند. اولین کاربرد آن در علوم روانشناسی بود که به تدریج به سایر حوزه‌ها از علوم طبیعی و پزشکی تا علوم اقتصادی و اجتماعی راه یافت (Mohades, 2010). از این روش در حوزه‌های اقتصادی نیز استفاده فراوانی شده است (Mohades, 2010). روش PCA یک مدل مناسب برای انجام تحقیق است چون باعث کاهش نمونه‌برداری و هزینه‌های مورد نیاز در تجزیه و تحلیل داده‌ها شده و می‌توان با استفاده از این روش متغیرهای پیش‌بینی کننده

شده است (Mahmodi et al, 2010). در سال‌های اخیر منابع آب‌های زیرزمینی این دشت به شدت افت داشته و کاهش متوسط حدوداً ۸ متری را نشان می‌دهد. در شکل (۱) محدوده منطقه مورد مطالعه و حدود آبخوان نشان داده شده است.

طول شرقی نسبت به نصف‌النهار گرینویچ واقع شده است. مساحت این دشت ۶۲۴۰۰ هکتار و مساحت حوزه آن ۲۵۵۰ کیلومترمربع و ارتفاع آن از سطح دریا ۱۸۷۶ متر بوده و در میان شهرستان‌های سنندج و قروه در استان کردستان واقع



شکل ۱- منطقه مورد مطالعه و حدود آبخوان در دشت قوره- دهگلان کردستان

(رابطه ۱)

$$V = W_D D + W_R R + W_A A + W_S S + W_T T + W_I I + W_C C$$

در رابطه (۱): D, R, A, S, T, I, C لایه‌های مدل (رتبه هر پارامتر) و W_i وزن‌های متناظر با لایه‌ها و همچنین V بیانگر شاخص آسیب‌پذیری می‌باشند.

(D): برای هر یک از سال‌های آماری میانگین عمق سطح ایستابی برای هر یک از چاه‌های مشاهده‌ای و در حال بهره‌برداری موجود در دشت محاسبه گردید. نقشه عمق سطح ایستابی با استفاده از روش درون‌یابی IDW به توان ۲ تهیه شد. (McCoy and Johanson, 2002). همچنین بر اساس جدول (۲) رتبه‌بندی گردید. (R): یکی از عوامل اصلی انتقال مواد آلاینده به درون آبخوان زیرزمینی است. میزان تغذیه با استفاده از روش پیشنهادی پیسکوپو استخراج گردید (Piscopo, 2001). در این روش با استفاده از میزان بارش متوسط بلندمدت منطقه، قابلیت نفوذپذیری و شیب، شاخص تغذیه منطقه با توجه به جدول (۱) مشخص شد.

مدل DRASTIC

استخراج نقشه‌های آسیب‌پذیری با توجه به منابع در دسترس و داده‌های موجود از روش‌های مختلفی صورت می‌پذیرد. به‌طور کلی این روش‌ها شامل: شاخص‌های همپوشانی، پردازش داده‌ها و روش‌های آماری است؛ که روش شاخص همپوشانی از اشتراک و تلفیق نقشه‌های متمایز مربوط به ویژگی‌های ژئوهیدرولوژیکی و جغرافیایی منطقه حاصل می‌شود. یکی از بهترین مدل‌های معرف در شاخص همپوشانی، مدل DRASTIC است که توسط EPA ارائه شده است. در این مدل میزان آسیب‌پذیری ارائه نمی‌گردد، بلکه یک شاخص بی‌بعد از آن را به دست می‌دهد. (Atiqur, 2008) مدل DRASTIC به‌صورت یک سرواژه از نام لایه‌های متمایزی به اسامی عمق تا سطح ایستابی، تغذیه خالص به آبخوان، زمین‌شناسی ناحیه اشباع، قشر خاک سطحی، شیب، اثر ناحیه غیراشباع و هدایت هیدرولیکی آبخوان تشکیل شده است؛ که حاصل جمع ضرب تک‌تک این لایه‌ها در ضرایب وزنی مربوط به آن لایه است و از رابطه (۱) محاسبه می‌شود. (Aller et al, 1987).

جدول ۱- تخمین شاخص تغذیه خالص آبخوان بر طبق روش پیسکوپو (۲۰۰۱)

میزان تغذیه		نفوذپذیری خاک		باراندگی		شیب	
رتبه	محدوده	فاکتور	محدوده	فاکتور	بارش (mm)	فاکتور	شیب (%)
۱	۳-۵	۱	خیلی کم	۱	۵۰۰ >	۱	۱۸ <
۳	۵-۷	۲	کم	۲	۵۰۰-۷۰۰	۲	۱۲-۱۸
۵	۷-۹	۳	متوسط	۳	۷۰۰-۸۵۰	۳	۶-۱۲
۸	۹-۱۱	۴	متوسط تا زیاد	۴	۸۵۰ <	۴	۲-۶
۱۰	۱۱-۱۴	۵	زیاد			۵	۲-۰

منطقه‌ای کردستان تهیه شد و بر طبق زمین‌شناسی و بافت‌شناسی منطقه، نقشه ناحیه غیراشباع تهیه و بر اساس جدول (۲) رتبه‌بندی گردید. (C): نقشه مربوط به قابلیت انتقال با استفاده از ضرایب هیدرودینامیکی به دست آمده از آزمایش‌های پمپاژ انجام شده در منطقه تهیه شد. سپس به وسیله لوگ چاه‌ها ضخامت آبرفت تعیین گردید. نقشه هم عمق نیز با استفاده از متوسط عمق ایستابی سالیانه تهیه شد و نقشه هدایت هیدرولیکی با استفاده از رابطه (۲) (Aller et al., 1987) و به کمک ابزار Raster Calculator در نرم‌افزار ARC-GIS تهیه گردید و بر اساس جدول (۲) رتبه‌بندی شد.

$$C = \frac{T}{B - D} \quad (\text{رابطه ۲})$$

در رابطه (۲): C: هدایت هیدرولیکی (m/day), T: قابلیت انتقال آبخوان (m²/day), B: ضخامت آبرفت (m), D: عمق سطح ایستابی (m)

(A): وضعیت زمین‌شناسی منطقه از طریق نقشه زمین‌شناسی سازمان زمین‌شناسی و اکتشاف معدنی کشور تهیه و بر اساس شاخص‌های آسیب‌پذیری موجود در جدول (۲) رتبه‌بندی گردید. (S): با استفاده از ۳۱ نمونه از خاک‌های آنالیز شده که از مرکز تحقیقات کشاورزی استان کردستان تهیه شد. برای دشت دهگلان بافت هر یک از نمونه‌ها تعیین و نقشه خاکشناسی تهیه و بر اساس جدول (۲) رتبه‌بندی گردید. (T): توپوگرافی به صورت شیب که تغییرات آن عامل کنترل‌کننده‌ای در نفوذ آلاینده‌ها به شمار می‌رود. هرچقدر شیب سطح زمین کمتر باشد زمان ماندگاری آلاینده‌ها روی سطح زمین طولانی‌تر بوده و پتانسیل آلودگی آبخوان زیرزمینی افزایش می‌یابد. مدل رقومی ارتفاعی منطقه (DEM) با استفاده از داده‌های ترازبایی شده توسط سازمان آب منطقه‌ای کردستان برای دشت قروه-دهگلان تهیه و سپس با استفاده از آنالیزهای مکانی نقشه شیب تولید و بر اساس جدول (۲) رتبه‌بندی گردید. (I): این لایه نقشه بر اساس تعداد ۳۰ لوگ چاه حفاری شده از سازمان آب

جدول ۲- ضرایب مدل دراستیک بر طبق پیشنهاد آلر و همکاران (۱۹۸۷)

وزن	A	رتبه	I	رتبه	T (%)	رتبه	D (m)
۶	رس شنی و کارست	۸	قلوه‌سنگ ریز با میان لایه شنی	۱	< ۱۸	۱	< ۴۵
۲	رس	۲	رس	۳	۱۲-۱۸	۲	۴۰-۴۵
۶	سنگ‌رس و قلوه	۴	رس شنی	۵	۶-۱۲	۳	۳۰-۴۰
۴	رس همراه با قلوه‌سنگ	۳	رس همراه با شن	۹	۲-۶	۴	۲۵-۳۰
۶	رس شنی همراه با قلوه‌سنگ	۲	رس همراه با آهک	۱۰	۰-۲	۵	۱۵-۲۵
۴	لوم رسی شنی	۵	رسی شنی و قلوه‌سنگ	رتبه	S	۷	۱۰-۱۵
۵	رس شنی	۷	شن و ماسه، قلوه‌سنگ و رس	۳	لوم رسی	۹	۰-۱۰
۸	شن و ماسه همراه با رس	۵	لوم رسی سیلتی همراه با گراول	۵	لوم	رتبه	(m/day) C
۶	سیلت و گراول			۴	لوم رس شنی	۱	۰-۵
				۶	لوم شنی	۲	۵-۱۵
						۴	۱۵-۳۵
						۶	۳۵-۵۰
						۸	۵۰<

(۲۰۱۰). (رابطه ۳) با استفاده از این روش می‌توان مدل‌های خطی را که ترکیب خطی از متغیرهای ورودی به مدل می‌باشند را مدل‌سازی نمود. در این تحقیق مدل آسیب‌پذیری آبخوان انتخاب گردید که تابعی خطی از پارامترهایی است که در بالا به آن‌ها اشاره شد. به دلیل داده مبنا بودن این روش جهت اعمال آن برای کاهش ابعاد مدلی مانند DRASTIC لازم است قبل از آن آزمون‌های لازم برای داشتن شروط کافی جهت اعمال PCA و استخراج مؤلفه‌های اساسی، کنترل و انجام شود. از آزمون‌های رایج برای این کار می‌توان آزمون Kaiser-Meyer-Olkin (KMO) و آزمون Bartlett's Sphericity نام برد.

تحلیل مؤلفه‌های اساسی (PCA)

تجزیه و تحلیل مؤلفه‌های اصلی (PCA) نوعی از تجزیه و تحلیل آماری است که تعداد کمتری از عوامل را به نام مؤلفه‌های اصلی از میان عوامل اولیه گزینش می‌کند، به طوری که تعدادی از اطلاعات کم‌اهمیت حذف می‌شود. در این روش متغیرهای موجود در یک فضای چند حالت همبسته به یک مجموعه از مؤلفه‌های غیرهمبسته به دست آمده مؤلفه‌های اساسی (PC) نامیده می‌شوند که از بردارهای ویژه ماتریس کوواریانس یا ماتریس همبستگی متغیر اصلی به دست می‌آیند (Mohades,)

مؤلفه‌ها گویند؛ که واریانس مؤلفه اصلی را می‌توان از رابطه (۸) به دست آورد.

$$Var(PC_i) = (EigenValue / k) \times 100 \quad (\text{رابطه ۸})$$

در رابطه (۸)، k تعداد متغیرها است. EigenValue مقدار ویژه مؤلفه i ام و Var_{PC} واریانس قابل توجیه در مؤلفه i ام

خوشه‌بندی

از آنجاکه تحلیل داده‌های خروجی از PCA برای مدل DRASTIC برای برنامه‌ریزان و مدیران منابع آب که نیاز به ابزارهای تصمیم‌گیری دارند، به دلیل هم مقیاس نبودن با کلاسه‌بندی‌های توصیه‌شده قابل‌اعمال نمی‌باشند، لازم است از ابزارها و متدهایی استفاده گردد تا نواحی مستعد آلودگی آبخوان تعیین و تصمیم‌های لازم جهت تحدید و اجرای پروژه‌های حفاظتی اقدام نمایند. برای انجام این کار روش‌های خوشه‌بندی به‌خوبی می‌توانند بر طبق وضعیت کیفی منطقه و شرایط آسیب‌پذیری آن پهنه‌های مختلف مدیریتی را در رابطه با استعداد آبخوان برای نشت و انتقال آلودگی تعیین کنند. تحلیل خوشه‌ای به دنبال سازمان‌دهی مجموعه‌ای از داده‌ها در یک سری خوشه است، به‌طوری‌که داده‌ها (که معمولاً هرکدام نشان‌دهنده برداری از مقادیر کمی در یک فضای چندین بعدی است) در هر خوشه بالاترین درجه شباهت را دارا بوده و با داده‌های موجود در خوشه‌های مختلف دارای حداکثر درجه عدم شباهت هستند. هدف خوشه‌بندی دستیابی سریع و مطمئن به اطلاعات همبسته و شناسایی ارتباط منطقی میان آن‌هاست؛ بنابراین، الگوریتم خوشه‌بندی می‌تواند در بسیاری از حوزه‌های موضوعی به کار گرفته شود. خوشه‌بندی، یکی از راهکارهای مؤثر در سازمان‌دهی اطلاعات به شمار می‌آید. با انجام فرایندهای خوشه‌بندی، حیطه گسترده‌ای از داده‌های پراکنده در گروه‌های مدون و سازمان‌یافته قرار می‌گیرند. الگوریتم‌های زیادی برای تجزیه خوشه‌ای پیشنهاد شده است. سه روش اصلی خوشه‌بندی که در علم آمار بسط و گسترش داده شده‌اند عبارت‌اند از: سلسله مراتبی، غیر سلسله مراتبی و فازی (Babaei *et al.*, 2005). در این مقاله از خوشه‌بندی سلسله مراتبی (رابطه ۹) و فازی (رابطه ۱۰) استفاده شده است. در پایان با استفاده از یکی از روش‌های صحت‌سنجی خوشه‌بندی، بهترین روش خوشه‌بندی انتخاب می‌شود.

$$\arg \min \left\{ \sum_{i=1}^n \sum_{j=1}^c \text{norm}(2, X_1, X_2) \right\} \quad (\text{رابطه ۹})$$

$$\arg \min \left\{ \sum_{i=1}^n \sum_{j=1}^c U_i^j \text{norm}(2, X_1, X_2) \right\} \quad (\text{رابطه ۱۰})$$

$$PC_i = E_i^T X = \sum_{j=1}^k e_{ij} X_k \quad (\text{رابطه ۳})$$

$$k = 1, 2, 3, 4, 5, 6, 7$$

در صورتی‌که آماره تست KMO (رابطه ۴) مربوط به این روش کمتر از ۰/۵ باشد، داده‌ها برای تجزیه و تحلیل عوامل اصلی مناسب نخواهند بود و اگر مقدار آن بین ۰/۵ تا ۰/۶۹ باشد باید با احتیاط بیشتر به تجزیه و تحلیل عوامل پرداخت؛ اما در صورتی‌که مقدار آن بزرگ‌تر از ۰/۷ باشد، همبستگی‌های موجود در بین داده‌ها برای تجزیه و تحلیل مناسب خواهد بود (Hutcheson and Nick, 1999).

در رابطه (۴): r_{ij}^2 بیانگر ضریب همبستگی بین متغیرهای i و j و a_{ij}^2 بیانگر ضریب همبستگی جزئی بین متغیرهای i و j است.

$$KMO = \frac{\sum_i \sum_{j \neq i} r_{ij}^2}{\sum_i \sum_{j \neq i} r_{ij}^2 + \sum_i \sum_{i \neq j} a_{ij}^2} \quad (\text{رابطه ۴})$$

با استفاده از تست Bartlett's Sphericity می‌توان در یک سطح اطمینان معین معنادار بودن همبستگی مقاطع همه متغیرها را باهم دیگر بررسی کرد. آماره این آزمون توزیع کای دو^۱ را با استفاده از رابطه (۵) الی (۷) تخمین می‌زند.

$$X^2 = ((N - k) \ln(S_p^2) - \sum_{i=1}^k (n_i - 1) \ln(S_i^2)) / (1 + \frac{1}{3(k-1)} (\sum_{i=1}^k (\frac{1}{n_i - 1}) - \frac{1}{N - k}))$$

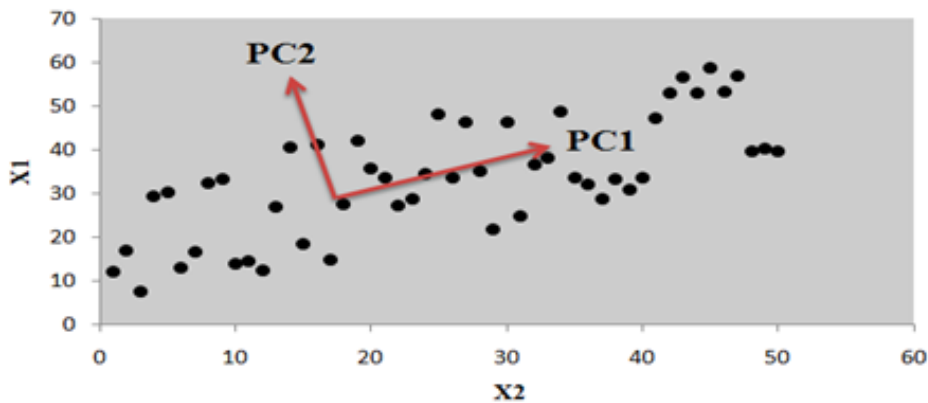
$$N = \sum_{i=1}^k n_i \quad (\text{رابطه ۶})$$

$$S_p^2 = \frac{1}{N - k} \sum (n_i - 1) S_i^2 \quad (\text{رابطه ۷})$$

در رابطه (۵) الی (۷) n_i متغیر i ام و S_i^2 واریانس مربوط به متغیر i ام است.

در این روش در صورتی‌که متغیرهای ورودی مدل را در مقابل همدیگر ترسیم کنیم، می‌توان راستاهای غالبی که واریانس داده‌ها در آن‌ها دیده می‌شوند را معین نمود. به هر یک از این راستاها که بر همدیگر متعامد می‌باشند مؤلفه یا مؤلفه‌های اصلی گویند. مثلاً برای یک مدل که ورودی‌های آن دو متغیر است و تصور آن ساده و قابل نشان دادن است مؤلفه‌های اصلی آن را می‌توان مطابق شکل (۲) نشان داد؛ که محورهای افقی و عمودی بیانگر متغیرها و محورهایی که نشان‌دهنده راستاها با واریانس‌های غالب مدل هستند را محور

1. Chie-Square



شکل ۲ - ابر متغیرهای ورودی و تصویر مؤلفه‌های اصلی آن (ویکی‌پدیا)

نتایج و بحث

مدل DRASTIC

پس از ایجاد لایه‌های مدل آسیب‌پذیری که در بالا به آن‌ها اشاره گردید، کلاسه‌ها برای هر یک از لایه‌ها و در همه نقاط مطالعاتی در سطح آبخوان استخراج شد و متغیرهای مورد استفاده در مدل جهت تعیین مؤلفه‌های اساسی توجیه‌کننده واریانس مدل، به‌عنوان ماتریس ورودی به PCA مورد استفاده قرار گرفت.

تحلیل مؤلفه‌های اساسی و نتایج آن

همان‌طوری که در بالا اشاره شد، قبل از انجام PCA بر روی داده‌ها باید از همبستگی لازم بین متغیرها برای انجام تجزیه و تحلیل اطمینان حاصل کنیم. برای بررسی این موضوع از دو آزمون Bartlett's Sphericity و KMO استفاده شد که نتایج آن در جدول (۳) برای Bartlett's Sphericity و جدول (۴) برای KMO آورده شده است.

با توجه به جدول (۳) مقدار p-value کوچک‌تر از ۰/۰۵ است بنابراین می‌توان گفت که بر اساس این تست همبستگی بین متغیرها معنی‌دار است و آبر آبر^۱ متشکل از تقاطع متغیرهای ورودی، غیره کروی بوده و در راستای یک مؤلفه اصلی که همان مؤلفه اول (PC₁) است توزین شده‌اند.

با توجه به جدول (۴) مقدار KMO بیشتر از ۰/۷ است؛ بنابراین همبستگی بین متغیرها تأیید شده و می‌توان از PCA بر روی این داده‌ها استفاده نمود. پس از انجام تست‌های فوق مؤلفه‌های اساسی متغیرهای مدل استخراج گردیده و نتایج آن در جدول (۵) دیده می‌شود. در شکل (۴) نیز مقادیر مربوط در سطور دوم و سوم جدول (۵) بر روی یک شکل ترسیم شده‌اند.

آزمون صحت‌سنجی خوشه‌بندی

نتایج حاصل از اعمال الگوریتم‌های خوشه‌بندی روی یک مجموعه داده، با توجه به روش‌های خوشه‌بندی می‌تواند بسیار متفاوت از یکدیگر باشد. شاخص‌های صحت‌سنجی برای سنجش میزان صحت نتایج خوشه‌بندی به‌منظور مقایسه بین روش‌های خوشه‌بندی مختلف یا مقایسه نتایج حاصل از یک روش با پارامترهای مختلف مورد استفاده قرار می‌گیرد؛ که به برخی از این شاخص‌ها در زیر اشاره شده است. همچنین دو معیار پایه اندازه‌گیری برای انتخاب خوشه‌های بهینه وجود دارد: (Kovacs et al, 2005)

الف- تراکم: داده‌های متعلق به یک خوشه باید تا حد ممکن به یکدیگر نزدیک باشد. معیار رایج برای تعیین میزان تراکم داده‌ها واریانس داده‌ها است.

ب- جدایی: خوشه‌ها خود باید به اندازه کافی از هم فاصله داشته باشد.

در این مقاله شاخص موردنظر برای صحت‌سنجی، شاخص دون انتخاب شد؛ که روابط (۱۱) الی (۱۳) برای محاسبه این شاخص است.

$$D = \min_{i=1 \dots C} \left\{ \min_{j=i+1 \dots C} \left(\frac{\text{norm}(2, C_i, C_j)}{\min_{k=1 \dots c} (\text{diam}(C_k))} \right) \right\} \quad (\text{رابطه ۱۱})$$

(رابطه ۱۲)

$$\text{norm}(2, C_i, C_j) = \min_{x \in C_i, y \in C_j} \{ \text{norm}(2, X_1, X_2) \}$$

$$\text{diam}(C_i) = \max_{x \in C_i, y \in C_j} \{ \text{norm}(2, X_1, X_2) \} \quad (\text{رابطه ۱۳})$$

در روابط (۱۱) الی (۱۳): n_c برابر تعداد خوشه‌ها و $d(x, y)$ بیانگر فاصله بین دو داده، همچنین C_i و C_j به ترتیب بیانگر i امین خوشه و j امین خوشه می‌باشند.

⁵ Hyper Cloud

جدول ۴- آزمون KMO برای کنترل و تعیین کفایت کمی داده‌ها در سطح ۵ درصد

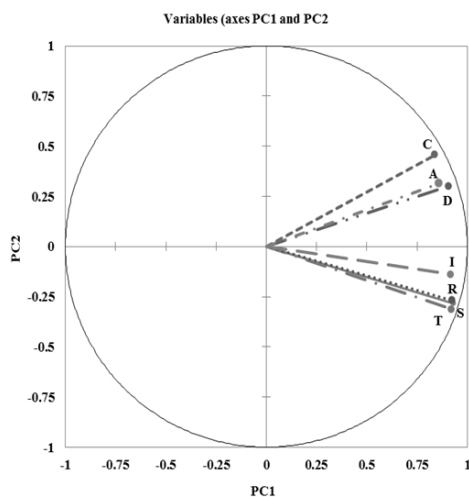
درصد	
D	۰/۸۴۸
C	۰/۸۰۰
A	۰/۹۱۸
KMO	۰/۸۴۲
T	۰/۷۸۷
S	۰/۸۰۲
R	۰/۸۸۴
I	۰/۸۷۸

جدول ۳- آزمون Bartlett's Sphericity برای وجوب انجام PCA در سطح ۵ درصد

درصد	
Chi-square (مقادیر مشاهداتی)	۳۸۹/۶۶۵
Chi-square (مقادیر محاسبه شده)	۳۲/۶۷۱
DF	۲۱
P-value	<۰/۰۰۰۱
Alpha	۰/۰۵

جدول ۵- مقادیر ویژه و واریانس قابل توجیه مربوط به هر یک از مؤلفه‌های اساسی مدل

	PC ₁	PC ₂	PC ₃	PC ₄	PC ₅	PC ₆	PC ₇
Eigenvalue	۵۱۰۷۶	۱/۴۴۴	۰/۲۳۵	۰/۱۲۶	۰/۰۵۷	۰/۰۴۸	۰/۰۱۴
Variability (%)	۷۲/۵۱۴	۲۰/۶۲۹	۳/۳۵۷	۱/۸۰۰	۰/۸۱۴	۰/۶۸۶	۰/۲۰۰
Cumulative (%)	۷۲/۵۱۴	۹۳/۱۴۳	۹۶/۵۰۰	۹۸/۳۰۰	۹۹/۱۱۴	۹۹/۸۰۰	۱۰۰/۰۰۰



شکل ۴- نمودار Scree، همبستگی مؤلفه‌های اول و دوم مدل با هر یک از متغیرها

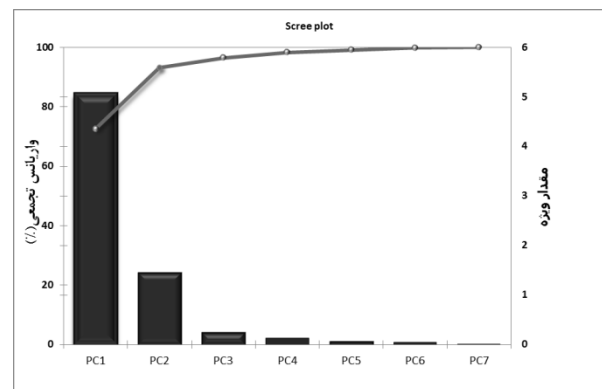
در شکل (۴) محور افقی معرف PC₁ و محور عمودی معرف PC₂ است که همبستگی متقاطع مؤلفه‌های اساسی اول و دوم (PC₁ و PC₂)، به دلیل توجیه بخش عمده واریانس ابر متشکل از متغیرها، ترسیم شده است. همان‌طور که مشخص است، بخش عمده‌ای از واریانس داده‌ها با استفاده از مؤلفه اول توجیه می‌گردد و این مؤلفه در همه متغیرها دارای بیشترین همبستگی در مقایسه با سایر مؤلفه‌ها است. همچنین از بین متغیرهای ورودی، متغیرهای I و D نسبت به سایر متغیرها همبستگی بیشتری را با PC₁ از خود نشان می‌دهند. همچنین مقادیر Score از رابطه (۱۴) به دست می‌آید.

(رابطه ۱۴)

$$Score_{PC_j} = \sum_{i=1}^m (ev_{i1}, ev_{i2}, \dots, ev_{im}) \cdot (x_{i1}, x_{i2}, \dots, x_{im})$$

در رابطه (۱۴) مقادیر ویژه و x ها متغیرها می‌باشند.

با توجه به جدول (۵) که مقدار واریانس نسبی و واریانس جمعی را برحسب درصد برای هر یک از مؤلفه‌های PCA نشان می‌دهد، مشخص است که مقدار واریانس برای PC₁ برابر با ۷۲/۵۱۴ و PC₂ برابر با ۲۰/۶۲۹ و مقدار واریانس سایر مؤلفه‌ها تقریباً ناچیز است. یکی از راه‌های یافتن تعداد مؤلفه‌های لازم برای انجام تحقیق مقدار واریانس است؛ یعنی مؤلفه‌هایی که درصد بیشتری از پراکندگی را توجیه می‌کنند، برای ادامه تحقیق کفایت می‌کنند و می‌توان از سایر مؤلفه‌ها صرف‌نظر نمود؛ بنابراین تنها PC₁ با بیشترین مقدار واریانس انتخاب شد و از سایر PC ها صرف‌نظر گردید. همچنین در شکل (۳) نمودار حاصل از نتایج به‌دست آمده، آورده شده است.



شکل ۳- نمودار واریانس جمعی و مقادیر ویژه مربوط به مؤلفه‌های اصلی مدل

در شکل (۳) مقادیر واریانس برای هر یک از مؤلفه‌ها نشان داده شده است که در آن محور افقی مؤلفه‌های اصلی و محور قائم سمت چپ واریانس و محور قائم سمت راست مقادیر ویژه را برای هر یک از این مؤلفه‌ها بیان می‌کند. واضح است که هر قدر از PC₁ به سمت PC₇ پیش می‌رویم از مقدار واریانس کاسته شده و در نهایت در PC₇ به صفر نزدیک می‌شود.

در این تحقیق از شش روش خوشه‌بندی Ward، Complete، Single، Average، Kmean، Cmeans بر روی داده‌های خروجی از PCA استفاده شد. نتایج حاصله از شش روش خوشه‌بندی در شکل (۶) نشان داده شده است.

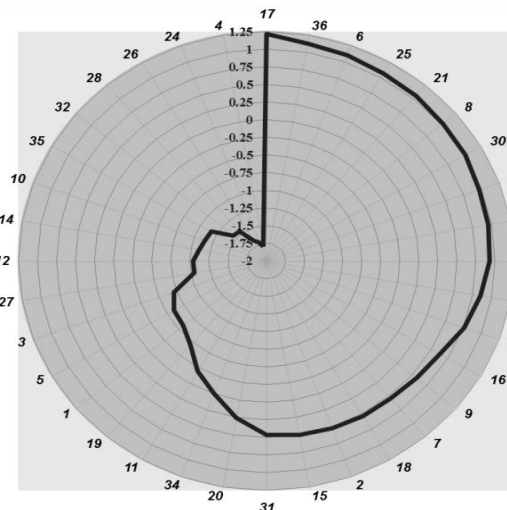
شکل (۶) نمودار حاصل از نتایج خروجی خوشه‌بندی به روش‌های Ward، Complete، Single، Average، Kmean، Cmeans است. اعداد بر روی محیط نمودار مربوط به شماره نقاط نمونه‌برداری و اعداد روی شعاع معرف Score محاسبه‌شده با مؤلفه اول است. همچنین رنگ‌های موجود در نمودار نشان‌دهنده محدوده هر کلاس حاصل از خوشه‌بندی است. همان‌گونه که در شکل نشان داده شده است، رنگ‌های سبز، آبی، زرد، قهوه‌ای و قرمز به ترتیب معرف محدوده کلاس‌های یک، دو، سه، چهار، پنج است. با استفاده از این نمودارها می‌توان حدود آسیب‌پذیری هریک از نقاط مشاهداتی را در خوشه‌های مربوط به آن‌ها دید همچنین با توجه به این نمودارها می‌توان دریافت که در هر یک از خوشه‌های محاسبه شده دارای چه حدودی بوده و کدام از نمونه‌های موردنظر در چه خوشه‌ای قرار گرفته‌اند.

به‌عنوان مثال تفسیر نمودار روش Ward در شکل (۶) به این صورت است. کلاس یک در روش Ward که محدوده آن با توجه به رنگ در نظر گرفته‌شده مشخص است (راهنمای بالای شکل‌ها)، بیانگر آسیب‌پذیری کم این محدوده بوده که نقاط (نمونه‌چاه‌های مشاهداتی) ۳، ۵، ۱۰، ۱۹، ۱۱، ۳۴، ۲۰ را در بر گرفته است. محدوده Score این کلاس بین ۰/۸- تا ۰/۳۵ قرار دارد. همچنین نقاط ۱۷، ۳۶، ۶، ۲۵، ۲۱، ۸، ۳۰، ۱۳، ۲۲، ۳۳، ۲۹، ۲۳ در کلاس ۴ قرار گرفته‌اند که در این روش نشان‌دهنده آسیب‌پذیری زیاد این نقاط است. محدوده Score این کلاس بین ۰/۶ تا ۱/۲۵ قرار دارد. وضعیت سایر کلاس‌ها در هر یک از روش‌های خوشه‌بندی روی نمودار مربوطه مشخص است که به‌طور خلاصه در جدول (۶) بیان شده است. در روش Single تعداد زیادی از نقاط در محدوده آسیب‌پذیری زیاد قرار گرفته است. ولی در سایر روش‌ها تعداد نقاط کمتری در محدوده آسیب‌پذیری زیاد قرار دارند. این موضوع می‌تواند نشان‌دهنده حساسیت بالای روش Single باشد. همچنین نتایج حاصل از روش‌های Ward، Average، complete شباهت زیادی با هم دارند. ولی روش Kmean با سایر روش‌های خوشه‌بندی تفاوت معناداری دارد.

نتایج حاصل از رابطه (۱۴) نرمال استاندارد نیست (Bazrafshan et al, 2014)؛ بنابراین باید مقادیر PC₁ تبدیل به نرمال استاندارد با میانگین صفر و انحراف معیار یک شود که یک معرف کمی برای شاخص دراستیک بوده و از مراحل پیش‌پردازش داده‌ها برای خوشه‌بندی است. رابطه (۱۵) برای استانداردسازی PC₁ مورد استفاده قرار می‌گیرد.

$$Z = (PC_1 - \overline{PC_1}) / SD \approx PC_1 / SD \quad (\text{رابطه ۱۵})$$

با توجه به اینکه مقدار میانگین PC₁ تقریباً نزدیک به صفر است. می‌توان از ترم میانگین PC₁ صرف‌نظر نمود. در رابطه (۱۵) $\overline{PC_1}$ بیانگر میانگین PC₁ ها و SD بیانگر انحراف معیار است (Bazrafshan et al, 2014). نتایج حاصل از PCA (ScorePC₁) نسبت به تمامی نقاط اندازه‌گیری شده در شکل (۵) نشان داده شده است.

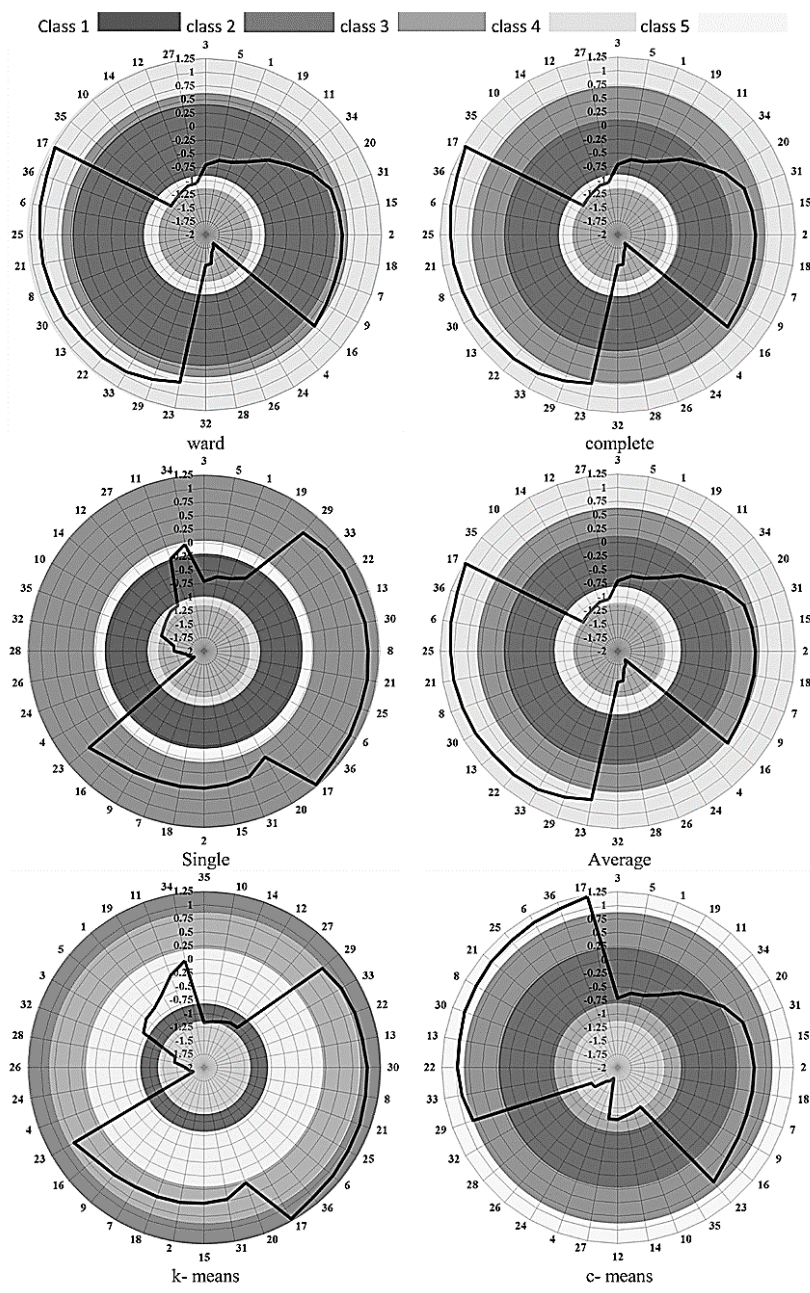


شکل ۵ - نمودار رادار Score برای مؤلفه اول به ازای کل نمونه‌های استفاده شده در مدل

شکل (۵) نتایج حاصل از PCA (ScorePC₁) نسبت به تمامی نقاط اندازه‌گیری شده است. همان‌طور که در شکل مشخص است تجزیه و تحلیل نتایج حاصله برای استفاده کاربردی از آن کمی مشکل است؛ بنابراین برای سهولت در استفاده از نتایج، از خوشه‌بندی استفاده می‌شود.

خوشه‌بندی و تعیین مناطق معرف وضعیت آسیب‌پذیری آبخوان

روش‌های زیادی برای انجام خوشه‌بندی وجود دارد و دقت هر یک از روش‌ها ممکن است در تحقیقات مختلف متفاوت باشد.



شکل ۶- نتایج حاصل از روش‌های مختلف خوشه‌بندی

جدول ۶- وضعیت هر کلاس در روش‌های خوشه‌بندی

Cluster	Ward	Complete	Single	Average	Kmean	Cmean
۱	آسیب پذیری کم	آسیب پذیری کم	آسیب پذیری کم	آسیب پذیری کم	آسیب پذیری خیلی کم	آسیب پذیری کم
۲	آسیب پذیری متوسط	آسیب پذیری متوسط	آسیب پذیری زیاد	آسیب پذیری متوسط	آسیب پذیری زیاد	آسیب پذیری متوسط
۳	آسیب پذیر نیست	آسیب پذیر نیست	آسیب پذیر نیست	آسیب پذیر نیست	آسیب پذیری متوسط	آسیب پذیری خیلی کم
۴	آسیب پذیری زیاد	آسیب پذیری زیاد	آسیب پذیری خیلی کم	آسیب پذیری زیاد	آسیب پذیر نیست	آسیب پذیر نیست
۵	آسیب پذیری خیلی کم	آسیب پذیری خیلی کم	آسیب پذیری متوسط	آسیب پذیری خیلی کم	آسیب پذیری کم	آسیب پذیری زیاد

در این مقاله باید نتایج یکی از این روش‌ها با بیشترین دقت انتخاب شود. برای این کار از آزمون صحت‌سنجی دان استفاده شد. نتایج حاصل از این آزمون در جدول (۸) آمده است.

جدول (۸) بیانگر نتایج حاصل از صحت‌سنجی خوشه‌بندی به روش دان بر روی همه روش‌های خوشه‌بندی است. همان‌طور که دیده می‌شود روش Single دارای بیشترین مقدار و روش‌های Kmean، Cmeans دارای کمترین مقدار می‌باشند. با توجه به اینکه مقدار ماکزیمم در روش دان شرط برتری روش است، از این‌رو روش Single به‌عنوان بهترین روش خوشه‌بندی در این تحقیق انتخاب شد.

جدول (۶) نشان‌دهنده وضعیت هر کلاس بر اساس نتایج حاصل از روش‌های خوشه‌بندی است. همان‌طور که مشخص است وضعیت کلاس‌ها در سه روش Ward, complete, Average شبیه به هم است. همچنین درصد محدوده هر کلاس در جدول (۷) ذکر شده است.

جدول (۷) بیانگر درصد کلاس‌ها در هر یک از روش‌های خوشه‌بندی است. همان‌طور که بیان شد بیش از ۵۰ درصد از سهم کلاس‌ها در روش Single به کلاس دو اختصاص یافته است که نشان‌دهنده آسیب‌پذیری زیاد نقاط موجود در این کلاس است. همچنین از بین روش‌های خوشه‌بندی استفاده‌شده

جدول ۷- سهم هر کلاس در روش‌های خوشه‌بندی برحسب درصد

Cluster	Ward	Complete	Single	Average	Kmean	Cmean
۱	۱۹/۴۴	۱۶/۶۷	۱۱/۱۱	۱۶/۶۷	۱۳/۸۹	۱۶/۶۷
۲	۱۹/۴۴	۲۲/۲۲	۵۵/۵۶	۲۲/۲۲	۳۰/۵۶	۲۵/۰۰
۳	۱۳/۸۹	۱۳/۸۹	۱۳/۸۹	۱۳/۸۹	۲۵/۰۰	۱۳/۸۹
۴	۳۳/۳۳	۳۳/۳۳	۱۳/۸۹	۳۳/۳۳	۱۳/۸۹	۱۳/۸۹
۵	۱۳/۸۹	۱۳/۸۹	۵/۵۶	۱۳/۸۹	۱۶/۶۷	۳۰/۵۶

جدول ۸- نتایج حاصل از آزمون صحت‌سنجی برای همه روش‌های خوشه‌بندی انجام‌شده

Ward	Complete	Single	Average	Kmean	Cmean
۰/۱۳۷۸۷۱۲	۰/۱۸۷۰۰۸۹	۰/۲۶۳۶۷۹	۰/۱۸۷۰۰۸۹	۰/۱۲۲۱۵۰۲	۰/۱۲۲۱۵۰۲

نتیجه‌گیری

از این‌رو در نمودار خوشه‌بندی Single میزان آسیب‌پذیری در نقاط مختلف دشت مشخص شده است. نتایج حاصل نشان می‌دهد که بیش از ۵۰ درصد از کل دشت در محدوده آسیب‌پذیری زیاد قرار دارد؛ بنابراین بایستی اقدامات لازم در نقاط آسیب‌پذیر در هنگام استفاده از این دشت انجام شود. مزیت‌های رویکرد بیان شده را می‌توان در قالب موارد زیر بیان کرد: ۱. برآورد مناسب و دقیق‌تر پارامترهای آسیب‌پذیری آبخوان. ۲. مقادیر آسیب‌پذیری این روش بی‌بعد و قابل مقایسه با سایر آبخوان دشت‌های دیگر است. ۳. در این روش میزان اثرگذاری و مشارکت هر یک از مؤلفه‌های مورد استفاده برای برآورد مقدار آسیب‌پذیری آبخوان قابل محاسبه است، به طوری که می‌توان مؤلفه‌های کم‌اثرتر را از روند محاسبات حذف و ابعاد و پیچیدگی مدل را کاست. ۴. با استفاده از روش‌های خوشه‌بندی و کنترل عملکرد آن‌ها، می‌توان نواحی در معرض خطر را مشخص و برنامه‌ریزی کاربری اراضی را مبتنی بر آسیب‌پذیری منطقه انجام داد.

در مطالعه صورت گرفته آسیب‌پذیری آبخوان با استفاده از روش DRASTIC و روش‌های تحلیل مؤلفه‌های اساسی (PCA) و تحلیل خوشه‌ای (CA) مدل‌سازی گردید. ابتدا متغیرهای روش DRASTIC آماده‌سازی و ایجاد گردید، سپس با استفاده از داده‌های این متغیرها روش تحلیل مؤلفه‌های اساسی جهت تقریب میزان آسیب‌پذیری آبخوان و برآورد یک ترکیب خطی از این متغیرها به کار گرفته شد. سپس مؤلفه اول به دلیل توجه بالای واریانس داده‌های مربوط به متغیرهای مدل، به‌عنوان مؤلفه معرف انتخاب گردید. برای تصمیم‌گیری در رابطه با مقادیر کمی برآورد شده با استفاده از مؤلفه اول از روش‌های تحلیل خوشه‌ای استفاده گردید و بهترین روش به‌وسیله شاخص‌های کارایی خوشه‌بندی برای خوشه‌سازی مقادیر آسیب‌پذیری معرفی گردید. با توجه به خوشه‌بندی انجام‌شده نقاط مختلف دشت از نظر آسیب‌پذیری به ۵ کلاس مختلف تفکیک شده است. از آنجا که روش خوشه‌بندی Single به‌عنوان بهترین روش با توجه به آزمون صحت‌سنجی دون انتخاب شد.

REFERENCES

- Ahmadi, F. Ranjbar, H. (2012). Dehghan Aquifer's water table declination assessment using GIS, 31th national geoscience. (In Farsi)
- Al - Adamat, R. Foster, I. and Baban, S. (2003).

- Groundwater vulnerability and risk mapping for the basaltic aquifer of the Azraq basin of Jordan using GIS, remote sensing and DRASTIC. *Applied Geography*, 23,303-324

- Aller, L. Benet, T. Lehr, J. Petty, R. and Hacket, G. (1987). DRASTIC: A standardized system for hydrogeologic settings (EPA- 60012-87-035) US EPA, Washington DC
- Atiqur, R. (2008). A GIS based DRASTIC model for assessing groundwatervulnerability in shallow aquifer in Aligarh, India. *Applied Geography*.
- Almasari, M.N. (2008). Assessment of intrinsic vulnerability to contamination for Gaza Coastal aquifer, Palestine
- Babiker, S. Mohamad, A. Hiyama, T. and Kato, K. (2005). A GIS- based DRASTIC model for assessing aquifer vulnerability in Kakamigahara Height, Gifu Prefecture, Central Japan. *Science of the Total Environment*, 345, 127-140
- Bazrafshan, J. Hejabi, S. and Rahimi, J. (2014). Drought monitoring using the multivariate standardized precipitation index (MSPD). *Water Resour Manage*. 28, 1045-1060.
- Babaei, Gholamreza, Feyzi, Avat. (2005). Classification of provinces based on population and health indicators by using fuzzy clustering statistical techniques, *Hakim journal*, No: 4, Tehran.
- Evans, B.M. and Myers, W.L. (1990). A CIS-based approach to evaluating regional groundwater pollution potential with DRASTIC. *Journal of Soil and Water Conservation*. 242-245.
- Fritch, T.G. Mcknight, C.L. Yelderman, J.C. and Arnold, J.G. (2000). An aquifer vulnerability assessment of the paluxy aquifer, Central Texas, USA, Using GIS and a modified DRASTIC approach. *Environment Management*, 25, 337-345.
- Hutcheson, G. and Nick, S. (1999). The multivariate social scientist: Introductory statistics using generalized linear models, thousand oaks, CA: sage publications.
- Khodaei, K. (2003). Intrinsic groundwater vulnerability mapping with concerning pollutants in Jovin plane using DRASTIC and GODS models. *Khorasan Water Resources Company*. (In Farsi)
- Khodaei, k. Shahsavari, B. and Etebari, B. (2007). Aquifer vulnerability assessment methods, GOD and DRASTIC, Iran. *J. Geol.* 4, 73-87. (In Farsi)
- Kovacs, F. Legany, C. and Babos, A. (2005). Clast validity measurement techniques. In *6th international symposium of Hungarian Researchers on Computational Intelligence*. Citeseer.
- McCoy, J. Johnston, K. (2002). Using ArcGIS Spatial Analyst. Esri Tutorial Help.
- Mohades, F. (2010). Principal component analysis method and factor analysis, a case study: dereviation of capital costs index and its effects on inflation. *Economics research series*, www.cbi.ir, 41. (In Farsi)
- Mahmodi, S. Fathi, P. Ebrahimi, M. (2010). Intelligent prediction of Dehgolan aquifer salinity using artificial neural networks. *4th national conference of water resources management*. (In Farsi)
- Nasir, M.F.M. Samsudin, M.S. Mohamad, I. Awaluddin, M.R.A. Mansour, M.A. Juahir, H. Ramli, N. (2011). River water quality modeling using combined principle component analysis (PCA) and multiple linear regressions (MLR): A case study at Klang River, Malaysia. *World Applied Sciencss Journal*, 73-82.
- National Academies Press. (1993). Ground Water Vulnerability Assessment.
- Piscopo, G. (2001). Groundwater vulnerability map, explanatory note, Castlreagh Catchment, NEW, Department of Land and Water Conservation. Australia.
- Rahman, A. (2008). A GIS based DRASTIC for assessing ground water vulnerability in shallow aquifer in Aligarh, India, *Applied Geography*, 28, 32-53.
- Secunda, S. Collin, M. and Mellou, A.J. (1998). Groundwater vulnerability assessment using a composite model combining DRASTIC with extensive land use. *Environmental Management*, 54, 39-57.
- Thareja, S. Choudhury, S. and Trivedi, P. (2011). Assessment of water quality of Ganga River in Kanpur by using principal components analysis, 2(5), 84-91.
- Wang, Y. wang, P. Bai, Y. Tian, Z. Li, J. Shao, X. Mustavich, L.F. Li, B.L. (2013). Assessment of surface water quality via multivariate statistical techniques: A case study of the Songhua River Harbin region, China, *Journal of Hydro-environment Reearch*, 7, 30-40.