

بهینه‌سازی تعداد چاه‌های نمونه‌برداری با رویکرد مکانی-زمانی در شبکه پایش سطح آب‌های زیرزمینی

(مطالعه موردی: دشت سراب)

نوید هوشنگی^{۱*}، میررضا غفاری رزین^۲

۱ و ۲. استادیار، گروه مهندسی نقشه‌برداری، دانشکده مهندسی علوم زمین، دانشگاه صنعتی اراک، اراک

(تاریخ دریافت ۱۴۰۰/۰۲/۰۲، تاریخ تصویب ۱۴۰۰/۰۵/۱۷)

چکیده

بهینه‌سازی چاه‌های نمونه‌برداری موجود در شبکه پایش آب‌های زیرزمینی به لحاظ کاهش هزینه نگهداری و بهبود کارایی و افزایش سرعت به‌روزرسانی اهمیت زیادی دارد. برای تعیین چاه‌های بااهمیت، عموماً از روش‌های زمین‌آماری (مانند کریجینگ) و یا تحلیل مؤلفه‌های اصلی (PCA) استفاده می‌شود. در روش‌های زمین‌آماری، بهینه‌سازی با توجه به موقعیت مکانی نمونه‌ها انجام می‌شود و اطلاعات زمانی چاه‌ها لحاظ نمی‌شود. در روش PCA تعیین چاه‌های شاخص با لحاظ اطلاعات زمانی چاه‌های موجود در همسایگی ایستگاه‌های پایش انجام می‌شود. در این تحقیق رویکردی بر اساس تلفیق این دو روش برای در نظر گرفتن اطلاعات مکانی و زمانی چاه‌ها ارائه شده و از آن برای کاهش تعداد چاه‌های نمونه‌برداری در تخمین سطح آب‌های زیرزمینی آبخوان دشت سراب استفاده شده است. تحقیق حاضر با اخذ مشاهدات ۴۷ چاه مربوط به سفره اول دشت سراب در سال ۱۳۹۷ در سه مرحله اصلی انجام گرفته است: ۱- جست‌وجوی اکتشافی داده‌ها؛ ۲- تعیین اولویت چاه‌ها با لحاظ اطلاعات زمانی چاه‌های حاضر در همسایگی با روش PCA؛ ۳- بررسی میزان تغییرات دقت سطح با روش کریجینگ با فرض حذف شدن چاه‌های با اولویت زمانی کم. نتایج نشان داد ۹ چاه دشت سراب (۱۹ درصد) دارای ارزش نسبی کمتر از ۰/۳ هستند که با حذف این چاه‌ها و ارزیابی خطای RMSE در نقاط حذف‌شده، مقدار این نقاط با ۴۶ سانتی‌متر خطا از طریق اطلاعات مکانی همسایه‌ها قابل محاسبه است. بنابراین، چاه‌های حذف‌شده اطلاعات مکانی زیادی وارد شبکه نمی‌کنند و با حذف آنها، می‌توان با افزایش دقت اندازه‌گیری سطح آب در بقیه چاه‌های نمونه‌برداری و صرفه‌جویی در زمان و هزینه، به همان دقت سطح اولیه رسید.

کلیدواژگان: آب زیرزمینی، چاه‌های نمونه‌برداری، زمین‌آمار، تحلیل مؤلفه‌های اصلی، دشت سراب.

مقدمه

آب‌های زیرزمینی یکی از منابع آبی مهم برای تأمین تمام بخش‌های زندگی بشر شامل نیازهای خانگی، صنعتی و کشاورزی به‌ویژه در مناطق خشک و نیمه‌خشک هستند [۱]. در مدیریت منابع آب زیرزمینی نظارت پیوسته بر چاه‌ها و پایش تغییرات کمی و کیفی آب‌های زیرزمینی بسیار مهم است [۲]. پایش مستمر آب‌های زیرزمینی نیازمند استقرار بهینه مجموعه‌ای از ایستگاه‌های اندازه‌گیری است [۳]. هدف از طراحی شبکه چاه‌های نمونه‌برداری (پیزومتر)، پایش کیفیت و سطح آب زیرزمینی به منظور مدیریت کارا و بهره‌برداری مؤثر از آبخوان است [۱ و ۴] که خروجی آن در اندازه‌گیری سطح آب زیرزمینی، پایش آلودگی و پیش‌بینی میزان فرونشست زمین مؤثر است [۵ و ۶]. شیوه پراکندگی ایستگاه‌های اندازه‌گیری در منطقه رابطه مستقیم با دقت برآوردها و در نتیجه، دقت تصمیم‌گیری‌ها دارد [۷].

عوامل متعددی (زمانی و مکانی) بر تغییرات سطح آب زیرزمینی مؤثرند، از جمله خصوصیات هیدروپنوماتیکی آبخوان، میزان بارش، وضعیت توپوگرافی، موقعیت مکانی عوامل تغذیه و تخلیه [۸]، که موجب پیچیده و غیرخطی شدن سیستم آب‌های زیرزمینی می‌شوند [۹ و ۱۰]. این پیچیدگی در تخمین و پایش آب زیرزمینی تصمیم‌گیران را به افزایش تعداد چاه‌های پیزومتری برای مدل‌سازی بهتر منطقه سوق می‌دهد. از طرفی، سامانه پایش طراحی شده باید مدام ارزیابی و اصلاح شود تا اهداف نظارتی خود را حفظ کند [۱ و ۱۱]. بنابراین، این دو مورد موجب افزایش هزینه برداشت اطلاعات می‌شود. بهینه‌سازی شبکه پایش و تعیین نقاط مهم در نمونه‌برداری به لحاظ کاهش حجم نمونه‌ها، صرفه‌جویی در هزینه، زمان و افزایش دقت نمونه‌برداری بسیار مهم است [۲ و ۴]. طراحی بهینه شبکه پایش آب‌های زیرزمینی، گامی ضروری در تشریح سیستم آب‌های زیرزمینی است و می‌تواند اطلاعات کمی و کیفی لازم برای انجام پژوهش درباره گستره مطالعه شده را فراهم بیاورد [۳ و ۱۲]. در صورت تشخیص نمونه‌های مستقل و وابسته در شبکه پایش، می‌توان در نمونه‌گیری‌های بعدی از نمونه‌های وابسته صرف‌نظر کرد. یک سیستم بهینه طراحی شده، نماینده‌ای از کیفیت ناحیه مورد پایش است و انتخاب تعداد بهینه ناحیه پایش و توزیع مکانی آنها با لحاظ بحث صرفه

اقتصادی یک نیاز اساسی و ضروری در آب‌شناسی است [۱۳].

هدف اصلی در بهینه‌سازی شبکه پایش، حذف اطلاعات زائد و افزونگی داده به منظور کاهش فراوانی برداشت اطلاعات است [۲ و ۱۴]. در این زمینه، رویکردهای مختلفی از جمله زمین‌آمار، آنتروپی، خوشه‌بندی، روش‌های بهینه‌سازی و تحلیل‌های آماری مانند ضریب تغییرات، تغییرات واریانس و تحلیل مؤلفه‌های اصلی^۱ (PCA) برای تعیین چاه‌های بااهمیت و ارزیابی شبکه پایش استفاده شده است. در رویکرد زمین‌آمار که بر اساس موقعیت مکانی چاه‌ها، بهینه‌سازی صورت می‌گیرد عموماً از روش‌های درون‌یابی مانند وزن‌دهی معکوس فاصله و کریجینگ استفاده می‌شود. در این رویکرد تعداد چاه‌های حذف‌شده با استفاده از اطلاعات چاه‌های موجود در همسایگی چاه‌های حذف‌شده و همبستگی مکانی موجود بین داده‌ها، برآورد می‌شود. این روش‌ها با وجود لحاظ کردن موقعیت مکانی چاه‌ها، نمی‌توانند اطلاعات زمانی چاه‌ها^۲ را لحاظ کنند. در روش آنتروپی میزان بی‌نظمی که هر چاه در شبکه ایجاد می‌کند، سنجیده می‌شود [۱۱]. در این رویکرد از نظریه آنتروپی استفاده می‌شود، که بر اساس آن، وقایع با احتمال وقوع کم، اطلاعات بیشتری در اختیار تصمیم‌گیران قرار می‌دهند، بنابراین چاهی که بی‌نظمی زیادتر ایجاد کرده باشد، اهمیت بیشتری دارد [۹]. در روش آنتروپی اگرچه اطلاعات چاه‌ها به صورت یکجا مورد بررسی قرار می‌گیرد، اما از میزان همبستگی مکانی داده‌ها استفاده نمی‌شود و خروجی روش به تعداد دسته‌هایی که داده‌ها در آن کلاسه‌بندی می‌شوند، وابسته است [۱۵]. در رویکرد خوشه‌بندی، داده‌های مشابه به چندین کلاس (دسته) تقسیم شده و سپس، از هر کلاس نمونه‌ای که نماینده بقیه نقاط باشد، انتخاب می‌شود [۱۴]. در این رویکرد نتایج بهینه‌سازی به نظرات کارشناسی و شیوه دسته‌بندی وابسته است [۱۶]. در رویکرد بهینه‌سازی از روش شبکه عصبی مصنوعی [۱۷] و روش‌های ابتکاری و فراابتکاری مانند ازدحام ذرات چندهدفه [۱۳]، تیرید شبیه‌سازی شده^۳ (SA)، کلونی مورچه و ژنتیک [۱۰] استفاده می‌شود که در آنها تابع

1. Principal Component Analysis

۲. منظور داده‌های ثبت‌شده برای یک چاه در زمان‌های مختلف

3. Simulated Annealing

حذف کرد، از ضریب تغییرات^۲ (CV) بین سطح اولیه و سطح ثانویه استفاده می‌شود [۱۸]. استفاده از ضریب تغییرات در صورتی قابل قبول است که فرض شود ضریب تغییرات با حذف چاه‌های غیر مؤثر افزایش یابد [۹]. به‌طور کلی، استفاده از ضریب تغییرات بدون در نظر گرفتن میزان همبستگی مکانی چاه‌ها و بدون در نظر گرفتن نقش چاه در کل منطقه می‌تواند موجب کاهش کارایی این روش شود.

از روش‌های زمین‌آماري برای تحلیل پدیده‌هایی که ماهیت مکانی دارند و فرضیات قوانین آمار کلاسیک مانند «تصادفی بودن داده‌ها صادق نیست»، استفاده می‌شود. در این رویکرد روش‌های درون‌یابی مانند وزن‌دهی معکوس فاصله و کریجینگ برای ایجاد سطح پیوسته از چندین نمونه با موقعیت مکانی مشخص استفاده می‌شود [۳ و ۱۰]. در تعیین چاه‌های شاخص با روش زمین‌آماري کریجینگ، فرض می‌شود یک چاه از مجموعه چاه‌های موجود حذف شود، مقدار چاه حذف‌شده با توجه به میزان همبستگی مکانی بین داده‌ها، توسط چاه‌های همسایگی آن محاسبه می‌شود، اگر مقدار تخمینی دقت کافی را نداشته باشد، گفته می‌شود که چاه حذف‌شده اهمیت زیادی دارد [۵ و ۲۴]. این روش در تحقیقات مختلفی برای تعیین اهمیت چاه‌ها در بهینه‌سازی شبکه پایش سطح آب زیرزمینی دشت شهرکرد [۵]، شبکه پایش دریاچه وینیپگ، کانادا [۳] و دشت تبریز [۲۵] استفاده شده است. ورودی این روش متوسط اطلاعات برداشت‌شده در یک ماه است و به صورت هم‌زمان نمی‌تواند اطلاعات چندین ماه را در نظر بگیرد. از طرفی، حذف چاه‌ها در این روش مرحله‌به‌مرحله و یک‌به‌یک انجام شده و بعد از حذف هر چاه باید کیفیت سطح خروجی بررسی شود [۲۶]. بنابراین، ایراد اصلی این روش، لحاظ نکردن اطلاعات سری زمانی داده‌ها و زمان‌بر بودن فرایند اجرای آن است.

در تحقیقاتی مشابه برای لحاظ اطلاعات مکانی و زمانی چاه‌های برداشت آب زیرزمینی، Destandau و Zaiter (۲۰۲۰) روشی مبتنی بر تئوری بی‌زی^۳ و بافت مکانی^۴ با هدف به حداکثر رساندن ارزش اقتصادی اطلاعات ارائه دادند. روش پیشنهادی آنها برای تشخیص سریع آلودگی تصادفی

بهینه‌شونده عموماً از روش‌های خوشه‌بندی [۱۷]، زمین‌آماري و یا آنتروپی استفاده می‌کنند [۷ و ۱۰]. در رویکرد آماری مانند تحلیل مؤلفه‌های اصلی، آماره داده‌ها در یک سری زمانی، با لحاظ اطلاعات چاه‌های موجود در همسایگی داده‌ها و بدون لحاظ اطلاعات زمین‌آماري مانند میزان همبستگی مکانی داده‌ها بررسی و تحلیل می‌شود. در بین رویکردهای یادشده روش غالبی وجود ندارد و تصمیم‌گیری در خصوص شیوه بهینه‌سازی شبکه پایش و روش استفاده‌شده به ماهیت شبکه، منطقه مطالعاتی و کیفیت اطلاعات موجود بستگی دارد [۱، ۳ و ۱۶]. در این بین، روش زمین‌آماري مبتنی بر کریجینگ و روش تحلیل مؤلفه‌های اصلی به دلیل سادگی و مقیاس‌پذیری مناسب^۱ محبوبیت زیادی دارند [۴، ۷، ۹، ۱۴، ۱۸ و ۱۹]. اما همان‌گونه که بیان شد، ضعف‌هایی در شیوه لحاظ اطلاعات در آنها وجود دارد.

روش PCA برای کاهش ابعاد داده و آشکارسازی الگوهای ساده‌تر در درون مجموعه‌ای از متغیرها استفاده می‌شود [۲۰ و ۲۱]. در این روش تعداد کمتری از عوامل به نام مؤلفه‌های اصلی از میان عوامل اولیه گزینش می‌شوند [۲۲]. هدف PCA یافتن زیرمجموعه‌ای از واریانس‌ها برای پوشش فضای چندبعدی متغیرها است [۱۹]. در این روش متغیرها در مؤلفه‌ها قرار می‌گیرند، به‌طوری که از مؤلفه اول به مؤلفه‌های بعدی درصد واریانس کاهش می‌یابد [۱۸]. از این‌رو، متغیرهایی که در مؤلفه‌های اول قرار می‌گیرند، تأثیرگذارترین هستند [۹]. در روش PCA داده‌های برداشت‌شده برای چاه‌ها در یک بازه زمانی مشخص بررسی می‌شوند و در آخر، هر چاهی که بیشترین تأثیر در شکل‌گیری مؤلفه‌های اصلی را داشته باشد، به عنوان چاه بااهمیت انتخاب می‌شود [۱۸]. این روش در تعیین چاه‌های مهم در شبکه پایش کروم آبخوان دشت بیرجند [۹]، منطقه کشت و صنعت سلمان فارسی [۱]، دشت ارومیه [۲۲]، حوضه باچیلیون در ایتالیا [۱۶]، در ایجاد شبکه پایش حس‌گرهای شبکه گاز [۲۱]، حس‌گرهای پایش آلودگی هوا، شبکه پایش میزان آسیب سازه‌ها [۲۳] و نظارت بر شبکه آماری چندمتغیره برای تشخیص ناهنجاری [۱۹] استفاده شده است. در آخرین مرحله از این روش برای تعیین تعداد چاه‌هایی که می‌توان

۱. منظور قابل پیاده‌سازی در مناطق مطالعاتی با وسعت‌های مختلف

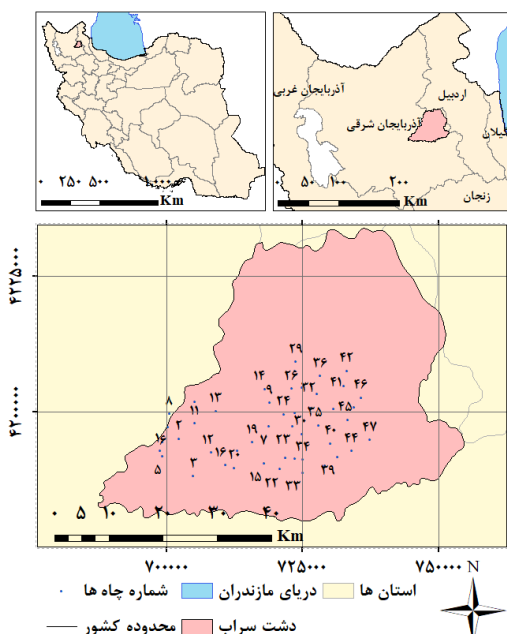
2. Coefficient of Variation
3. Bayesian
4. Geographical context

منطقه مطالعه شده و داده‌های استفاده شده

دشت سراب - دوزدوزان به وسعت ۱۱۳۴۰۰ هکتار در محور رودخانه آجی چای و در امتداد شرقی - غربی واقع شده است (شکل ۱). این دشت، شهر سراب را احاطه کرده و دارای کشتزارها و مراتع وسیع است. سفره آب‌های آزاد این دشت بیشتر از جریان آبی کوهستان‌های سبلان و بزقوش تأمین می‌شود [۲۷]. این دشت شش دره آبرفتی دارد که رسوبات آنها در قسمت پایین دست به هم ملحق می‌شوند و دشت سراب را به وجود می‌آورند و شامل دره کلیان، دره آغمیون، دره کرد امیر، دره زارلیک، دره شیر و انجیک و دره ینگجه می‌شوند. این دره‌ها در واقع محل‌های ورود عمده آب‌های سطحی و زیرزمینی دشت محسوب می‌شوند و رودخانه آجی چای در پایین‌ترین سطح دشت قرار گرفته است. جهانبخش اصل و همکاران (۱۳۸۸) خطر آسیب‌پذیری پوشش گیاهی در برابر خشک‌سالی‌های احتمالی در منطقه غرب دشت سراب را بسیار زیاد ارزیابی کرده‌اند [۲۷]. در سال ۱۳۷۰ معادل ۶۱/۴۲ درصد وسعت دشت سراب - دوزدوزان مورد استفاده زراعی قرار داشته است. ۶۵ هزار هکتار از اراضی منطقه به صورت دیم و ۵۱ هزار هکتار به صورت کشت آبی هستند. وجود ۱۷ سد کوچک در منطقه سراب با توجه به وسعت دشت سراب کافی نیست و سدها فقط ۳۷ میلیون مترمکعب آب را ذخیره می‌کنند، در حالی که بیشترین میزان

آب‌های زیرزمینی مناسب بوده و در آن وضعیت کلی شبکه برای پایش‌های سالیانه سطح آب مد نظر نبوده است [۲]، بنابراین قابلیت اجرایی آن در تخمین سطح آب زیرزمینی بررسی نشده است. همچنین، هوشنگی (۱۳۹۸) با استفاده از بهینه‌سازی کلونی مورچه و روش کریجینگ، چاه‌های مؤثر در یک سال را بررسی کرد و رویکردی برای انتخاب چاه‌های شاخص ارائه داد. در این رویکرد برای رفع مشکلات روش زمین‌آماری، ابتدا اطلاعات مکانی هر ماه به صورت جداگانه بررسی شد و در نهایت، مجموع دفعاتی که هر چاه به عنوان چاه مؤثر انتخاب شده، در تعیین شبکه بهینه استفاده شد [۲۵].

در دشت سراب بیشتر آب استفاده شده برای شرب و کشاورزی از منابع آب زیرزمینی تهیه می‌شود که موجب شده طی زمان از سطح آب‌های زیرزمینی کاسته شود و آبخوان این دشت را با بحران کم‌آبی روبه‌رو کند. این دشت در سال‌های اخیر از نظر استحصال جزء دشت‌های آزاد است و در همسایگی دشت ممنوعه اردبیل قرار دارد. افت سطح آب در سال‌های اخیر در این دشت، تصمیم‌گیران را به سمت پایش مستمر و تلاش برای پیشگیری از قرار گرفتن در وضعیت بحرانی سوق داده است. تا کنون طرحی نظام‌مند و علمی برای اصلاح و بهینه‌سازی شبکه پایش کیفی منابع آب زیرزمینی دشت با بیش از ۶۴ ایستگاه پایش بالقوه، ۴۷ چاه در سفره اول و ۱۷ چاه در سفره دوم ارائه نشده است. با لحاظ محدودیت‌ها و کمبودها در سطح آب زیرزمینی دشت سراب و نبود تحقیق در زمینه بهینه‌سازی شبکه پایش در این دشت، هدف اصلی این تحقیق، ارائه رویکردی نوین برای یافتن چاه‌های بااهمیت در پایش آب زیرزمینی دشت سراب با لحاظ اطلاعات مکانی-زمانی چاه‌ها است. به این منظور، بعد از بررسی کیفیت و صحت داده‌های سفره اول دشت سراب با استفاده از روش تحلیل مؤلفه‌های اصلی و روش زمین‌آماری کریجینگ، اولویت چاه‌ها با توجه به داده‌های زمانی و مکانی^۱ بررسی می‌شود. در تحقیقات پیشین عموماً از داده‌های زمانی چاه‌ها یا داده‌های مکانی چاه‌ها استفاده شده و استفاده هم‌زمان از دو روش PCA و زمین‌آماری و بهره‌گیری از مزایای دو روش بررسی نشده است.



شکل ۱. آبخوان دشت سراب و چاه‌های پی‌زومتری موجود

۱. منظور داده‌های ثبت شده برای چاه‌های کناری با توجه به موقعیت مکانی چاه

مرحله اول: جست‌وجوی اکتشافی داده‌ها

داده‌های برداشت‌شده در سری‌های زمانی متفاوت تحت تأثیر اشتباهات و خطاهای مختلفی هستند [۱۲]، بنابراین در ایستگاه‌های پایش عموماً آزمون‌های مختلفی برای تحلیل اکتشافی و بررسی کیفیت داده‌ها مانند بررسی شیوه توزیع، نوع روند، میزان وابستگی داده‌ها به هم و بررسی همگنی داده‌ها انجام می‌گیرد [۲۵]. شناسایی خطاها و داده‌های پرت^۱ برای ارزیابی و بازسازی داده‌ها، به آن سبب بااهمیت است که وجود این گونه داده‌ها می‌توانند نتایج پژوهش را به‌ویژه در پیش‌بینی مکانی و زمانی تحت تأثیر قرار داده و اعتبار پژوهش را دچار تردید کند. محیط‌های سیستم اطلاعات مکانی^۲ (GIS) در این زمینه ابزارهای مناسبی را فراهم آورده‌اند [۱۷].

در این مرحله ابتدا شیوه توزیع داده‌های هر ماه بررسی و داده‌های پرت شناسایی می‌شود. برای بررسی داده پرت از آزمون دیکسون استفاده می‌شود. این آزمون روشی برای یافتن نقاط دورافتاده یا پرت در مجموعه داده‌های کوچک است که به‌طور نرمال توزیع یا از یک جامعه نرمال گرفته شده‌اند. این نقاط ممکن است نسبت به بقیه مقادیر مجموعه داده‌ها، بسیار بزرگ‌تر یا بسیار کوچک‌تر باشند. داده‌های پرت شناسایی‌شده، بلافاصله حذف نمی‌شوند، بلکه نیاز به بررسی کارشناسی و دقت بیشتر دارند. در این مرحله همچنین چاه‌های مشترکی که طی ۱۲ ماه در سال ۱۳۹۷ فعال بوده‌اند، شناسایی و داده‌های جافتاده^۳ برای هر چاه محاسبه می‌شود. داده‌های جافتاده یکی از مشکلات رایج در داده‌های سری زمانی هستند [۱۲ و ۲۹]. برای محاسبه مقدار داده‌های جافتاده از تحلیل حداقل مربعات جزئی غیر تکراری غیرخطی^۴ (NIPALS) استفاده می‌شود. روش NIPALS تجزیه و تحلیل مؤلفه اصلی را با مقادیر جافتاده امکان‌پذیر می‌کند. این روش روی مجموعه داده‌ها اعمال شده و مدل به‌دست‌آمده از آن برای پیش‌بینی مقادیر جافتاده استفاده می‌شود [۳۰].

مرحله دوم: تحلیل مؤلفه‌های اصلی (PCA)

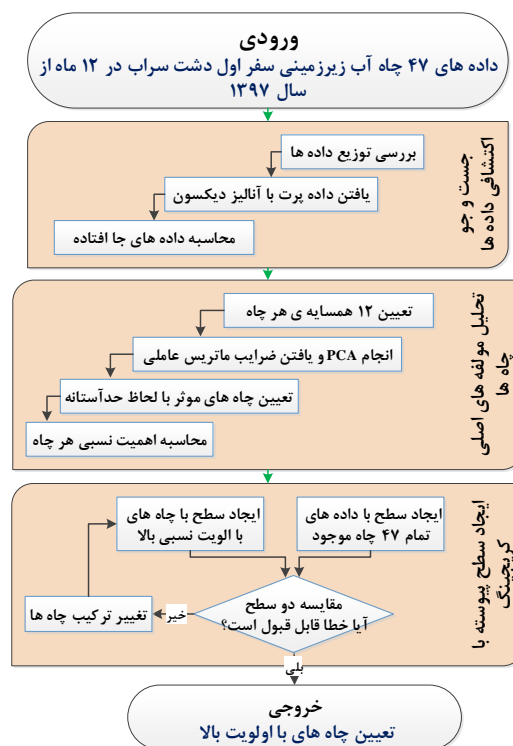
در تحلیل مؤلفه‌های اصلی چاهی که داده‌های آن دارای

تأمین منابع آبی در منطقه از طریق چاه‌ها و قنات با ۱۶۵ میلیون مترمکعب است. این در حالی است که دشت سراب با شوری خاک مواجه بوده که یکی از عوامل اصلی آن، شوری و کاهش آب‌های زیرزمینی بیان شده است [۲۸]. بنابراین، خشک‌سالی‌های احتمالی و پایش پیوسته سطح و کیفیت آب زیرزمینی طراحی و بهینه‌سازی شبکه نظارت آب زیرزمینی در دشت سراب را مهم کرده است.

در این تحقیق اطلاعات ۶۸ چاه پیزومتری آبخوان دشت سراب در ۱۲ ماه از سال ۱۳۹۷ تهیه شده و از ۴۹ داده سفره اول استفاده شد. بعد از تحلیل اولیه دو چاه به عنوان داده پرت شناسایی شد و در نهایت، از ۴۷ چاه نمونه‌برداری در فرایند کاهش چاه‌های پیزومتری استفاده شد. موقعیت و شماره چاه‌های نمونه‌برداری شده در شکل ۱ نمایش داده شده است. البته، تعداد داده‌های ثبت‌شده در هر ماه متفاوت بوده است. این اطلاعات از سامانه آمار و گزارش شرکت مدیریت منابع آب ایران دفتر مطالعات پایه منابع آب دریافت شده است.

روش شناسی

مراحل اجرای این تحقیق مطابق شکل ۲ در سه مرحله اصلی بیان شده که در ادامه هر یک از این مراحل تشریح می‌شود.



شکل ۲. مراحل اجرای تحقیق

1. Outlier
2. Geospatial Information Systems
3. Missing Data
4. Nonlinear Iterative Partial Least Squares

مرحله چهارم: در این مرحله اهمیت نسبی چاه‌ها با لحاظ نتایج ۱۲ ماه مشخص می‌شود. بنابراین، نتایج مرحله سوم جمع می‌شود و اهمیت نسبی هر چاه با استفاده از فراوانی مؤثر هر چاه به دست می‌آید. برای محاسبه اهمیت نسبی هر چاه مجموع دفعاتی که هر چاه در محاسبات چاه‌های دیگر به عنوان چاه مؤثر شناخته شده است، شمارش شده و به مجموع دفعاتی که هر چاه در همسایگی چاه‌های دیگر ظاهر شده، تقسیم می‌شود.

مرحله سوم: ایجاد سطح پیوسته با روش کریجینگ
در این مرحله با استفاده از خروجی مرحله قبل (اولویت نسبی چاه‌ها) مشخص می‌شود که چند چاه با اولویت نسبی کم را می‌توان حذف کرد، به گونه‌ای که در دقت سطح جدید ایجاد شده تغییر زیادی ایجاد نشود. در بیشتر مطالعات پیشین (مانند پژوهش خاشعی و همکاران [۹]) در این مرحله از متوسط ضریب تغییرات سطح برای تعیین تعداد نقاط حذف شده استفاده شده که بدون توجه به موقعیت و میزان همبستگی مکانی داده‌ها و پهنه ایجاد شده محاسبه شده است. توجه به این نکته می‌تواند تعیین کننده تعداد چاه‌های قابل حذف باشد که برای تصمیم‌گیری در مطالعات پدیده‌های جغرافیایی از داده‌های برداشت شده به صورت نقطه‌ای، عموماً یک سطح پیوسته با استفاده از روش‌های درون‌یابی ایجاد می‌شود [۵]. هر چه شبکه پایش طراحی شده بهینه باشد، سطح ایجاد شده به واقعیت زمینی نزدیک‌تر خواهد بود. بنابراین، حذف چاه‌های کم‌اهمیت نیز باید تا حدی انجام شود که در سطح پیوسته ایجاد شده بعد از حذف چاه‌ها نسبت به حالت اولیه، خطای فاحشی ایجاد نشود. این مرحله در واقع نوآوری اصلی این تحقیق است.

در این مرحله، میزان تغییرات سطح آب به ازای حذف چاه‌های با اولویت نسبی کم محاسبه می‌شود. بنابراین، ابتدا برای هر ماه با استفاده از روش‌های درون‌یابی و کل مجموعه نقاط موجود سطحی پیوسته ایجاد می‌شود. سپس، با حذف چاه‌های کم‌اهمیت سطح جدیدی ایجاد شده و میزان خطا در ایستگاه‌های حذف شده محاسبه می‌شود. اگر اختلاف مقدار برآورد شده از سطح جدید با مقادیر برداشت شده زمینی کم باشد، نشان می‌دهد مقدار ایستگاه پایش توسط ایستگاه‌های همسایه قابل اندازه‌گیری است و در نتیجه، حذف چاه‌های

همبستگی زیاد و معناداری با مؤلفه‌های اصلی باشد، به عنوان چاه مؤثر و مهم معرفی می‌شود [۱۴]. اجرای این روش برای تعیین چاه‌های مؤثر در چهار مرحله انجام می‌گیرد [۱ و ۱۸]:
مرحله اول: در این مرحله چاه‌های واقع در همسایگی هر چاه پیژومتری شناسایی می‌شوند. آنالیز PCA روی یک ماتریس اعمال می‌شود که در آن تعداد سطرها باید بزرگ‌تر یا مساوی تعداد ستون‌ها باشد. در این تحقیق ستون‌ها بیانگر چاه‌ها و سطرها بیانگر مشاهدات در سال‌ها یا ماه‌های متفاوت‌اند. بنابراین، تعداد چاه‌های واقع در همسایگی باید به گونه‌ای انتخاب شوند که تعداد آنها از تعداد زمان‌های برداشت اطلاعات بزرگ‌تر یا مساوی باشد [۲۲]. در این تحقیق از اطلاعات ۱۲ ماه چاه‌های پیژومتری دشت سراب استفاده می‌شود، بنابراین ستون‌های این ماتریس ۱۲ همسایه مجاور هر چاه خواهد بود.

مرحله دوم: در این مرحله آنالیز PCA به منظور یافتن عامل‌های اصلی و وابستگی آنها به داده‌های چاه‌های دخیل اجرا می‌شود. از جمله خروجی‌های آنالیز PCA بردارهای ویژه، نمودار درصد واریانس هر مؤلفه، رابطه بین هر چاه با مؤلفه‌های اصلی است. رابطه بین هر چاه با مؤلفه‌های اصلی از ضریب همبستگی بین مؤلفه‌های اصلی و داده‌های زمینی محاسبه شده و در ماتریس عاملی^۱ ارائه می‌شود [۱۸]. درایه‌های ماتریس عاملی سهم متغیرها را در مؤلفه‌ها نشان می‌دهد. هر چه این مقدار بیشتر باشد، همبستگی بین چاه و مؤلفه اصلی و در نتیجه، اهمیت چاه در شکل‌گیری مؤلفه‌های اصلی بیشتر است [۹ و ۲۰]. این مرحله به تعداد چاه‌های پیژومتری موجود تکرار می‌شود تا میزان اهمیت هر چاه تعیین شود.

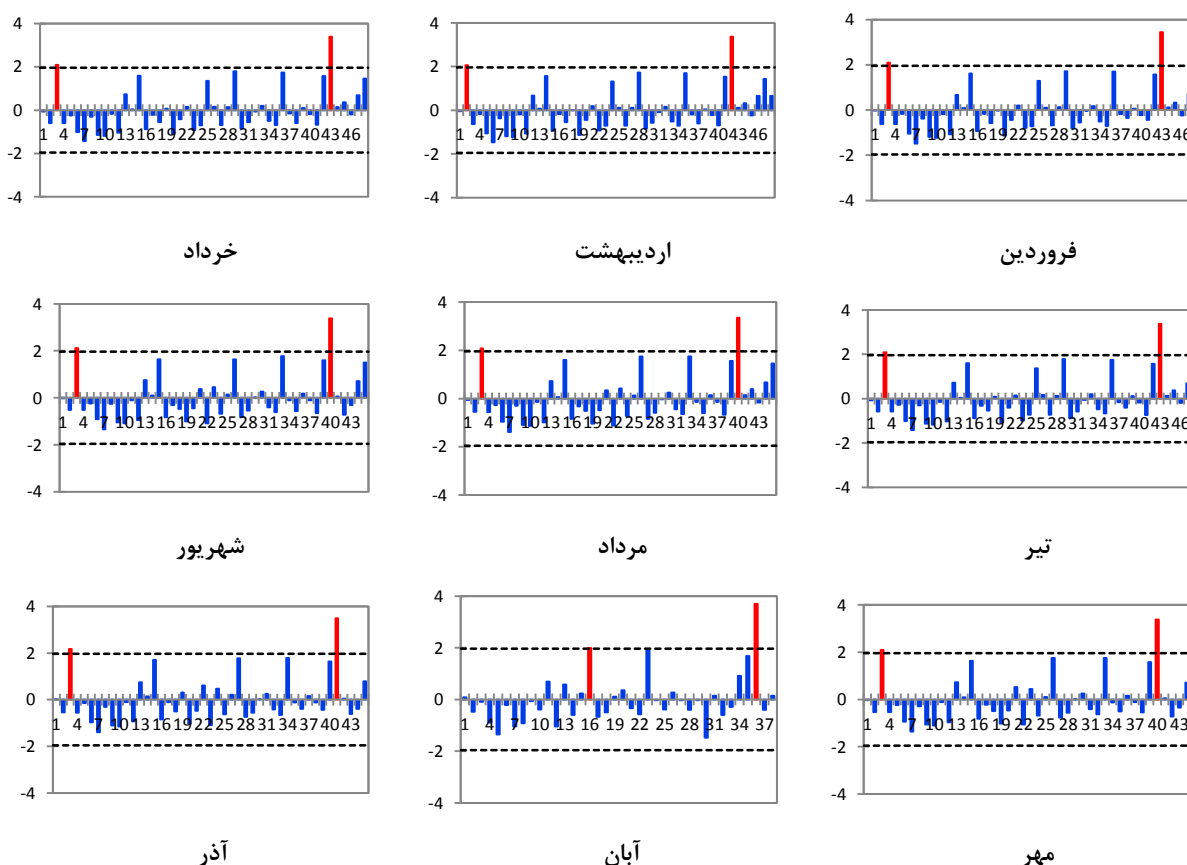
مرحله سوم: در این مرحله با توجه به حد آستانه لحاظ شده برای داریه‌های ماتریس عاملی (سهم هر چاه در مؤلفه‌های اصلی) اهمیت چاه‌های مورد نظر تعیین می‌شود. خاشعی سیوکی و همکاران (۱۴۰۰) از حد آستانه ۰/۷۵ [۹] و Wang و همکاران (۲۰۱۸) از حد آستانه ۰/۸۵ [۱۴] برای تعیین چاه‌های مؤثر استفاده کردند، اما در بیشتر تحقیقات پیشین از حد آستانه ۰/۹ استفاده شده است [۱، ۱۸، ۲۰ و ۲۲]. در این تحقیق، در هر ماتریس عاملی، چاهی که دارای همبستگی بیشتر از ۰/۹ با مؤلفه‌های اصلی باشد، به عنوان چاه مؤثر انتخاب می‌شود.

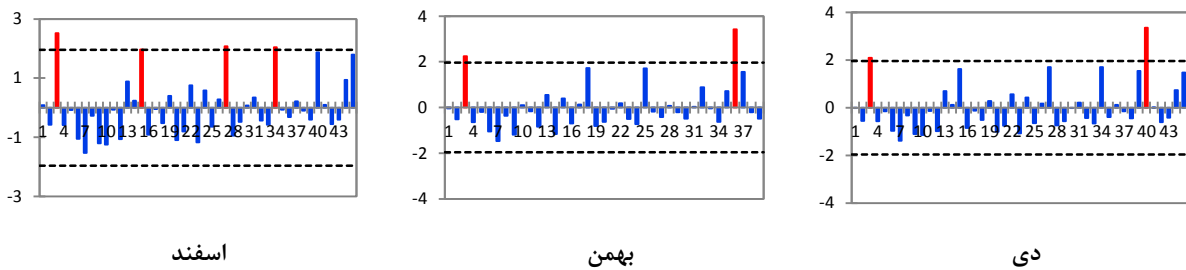
تجربی و واریانس تغییر است [۱۳]. در این مرحله برای محاسبه خطا در نقاط حذف‌شده از معیار خطای جذر میانگین مربعات (RMSE)^۱ در کریجینگ ساده استفاده می‌شود. از RMSE در تحقیقات مختلفی برای ارزیابی دقت مکانی شبکه پایش استفاده شده است [۱۰ و ۳۱].

نتایج و بحث

در این بخش خروجی محاسبات و تحلیل‌های انجام‌شده گنجانده شده و نتایج تحقیق بررسی می‌شوند. بعد از جمع‌آوری داده‌های مورد نیاز (۴۹ داده سفره اول آبخوان دشت سراب) آزمون تعیین داده‌های پرت با استفاده از تحلیل دیکسون در افزونه Xlstat در نرم‌افزار اکسل انجام شد. نتایج بررسی داده‌های پرت برای هر ماه در شکل ۳ نمایش داده شده است.

کم‌اهمیت تأیید می‌شود. در این مرحله به ازای حذف چاه‌ها با اولویت‌های نسبی مختلف، متوسط مقدار خطا در ماه‌ها محاسبه شده و منحنی خطا ترسیم می‌شود. محاسبه اولویت نسبی چاه‌ها بر اساس اطلاعات زمانی در مرحله قبل در واقع کمک می‌کند که سریع‌تر مجموعه ایستگاه‌های با ارزش مکانی تعیین شوند. در این مرحله از روش زمین‌آماری کریجینگ برای درون‌یابی و ایجاد سطح پیوسته استفاده می‌شود. روش کریجینگ جزء محبوب‌ترین روش‌های درون‌یابی زمین‌آماری است که در تحقیقات مختلفی برای پهنه‌بندی استفاده شده است [۳، ۵، ۷ و ۱۰]. این روش بهترین برآوردگر خطی نارایب است و پیش‌فرض نرمال بودن داده‌ها و وجود همبستگی مکانی بین داده‌ها را دارد [۳ و ۴]. کریجینگ متناسب با پیش‌فرض‌های اولیه انواع مختلفی از جمله ساده، معمولی، شاخص و عمومی دارد [۳۱]. کریجینگ ساده به عنوان حالت پایه کریجینگ شامل دو بخش نیم‌تغییرنمای





شکل ۳. تعیین داده‌های پرت با استفاده از تحلیل دیکسون

محاسبات اطلاعات این دو چاه حذف شدند. در بقیه موارد در ماه‌های آبان و اسفند بررسی کارشناسی نشان داد این داده‌ها برداشت واقعی زمینی هستند و به عنوان داده پرت نیستند. ویژگی‌های آماری داده‌های استفاده‌شده بعد از حذف داده‌های پرت در جدول ۱ مشاهده می‌شود. تعداد چاه‌های نمونه‌برداری در سفره اول دشت سراب، ۴۷ ایستگاه است که در هر ماه حداقل در یک چاه برداشت اطلاعات صورت نگرفته است.

تعداد داده‌های برداشت‌شده در هر ماه متفاوت است و در ماه‌های آبان و بهمن کمترین تعداد چاه نمونه‌برداری شده‌اند. این تعداد برای آبان‌ماه، ۳۸ اندازه‌گیری و برای بهمن‌ماه، ۳۹ اندازه‌گیری است. همان‌گونه که از شکل‌های یادشده مشخص است دو چاه در بیشتر ماه‌ها به عنوان چاه داده پرت معرفی شده است که بعد از بررسی کارشناسی مشخص شد که این دو چاه مربوط به سفره دوم دشت سراب هستند و به‌اشتباه در بین داده‌های سفر اول لحاظ شده‌اند که در ادامه

جدول ۱. اطلاعات آماری داده‌های استفاده‌شده (اعداد برحسب متر)

تعداد چاه	کمینه	بیشینه	چارک اول	میانه	چارک سوم	میانگین	انحراف معیار	
۴۶	۱۶۳۱/۳۳	۱۷۵۸/۱۶	۱۶۶۴/۷۱	۱۶۸۲/۰۰	۱۶۹۴/۸۶	۱۶۸۴/۴۸	۳۱/۸۲	فروردین
۴۵	۱۶۳۱/۵۵	۱۷۶۰/۳۳	۱۶۶۲/۸۰	۱۶۸۲/۱۵	۱۶۹۵/۳۶	۱۶۸۴/۱۴	۳۲/۴۳	اردیبهشت
۴۶	۱۶۳۱/۳۷	۱۷۶۲/۱۴	۱۶۶۲/۰۷	۱۶۸۰/۹۰	۱۶۹۵/۳۲	۱۶۸۳/۷۴	۳۲/۶۸	خرداد
۴۶	۱۶۳۱/۰۴	۱۷۶۱/۲۹	۱۶۶۰/۸۴	۱۶۷۹/۳۸	۱۶۹۴/۰۸	۱۶۸۳/۱۳	۳۲/۸۰	تیر
۴۳	۱۶۳۰/۸۳	۱۷۵۹/۷۸	۱۶۶۰/۸۱	۱۶۷۸/۶۹	۱۶۹۷/۹۶	۱۶۸۲/۸۶	۳۳/۲۷	مرداد
۴۳	۱۶۳۰/۷۰	۱۷۵۹/۲۰	۱۶۵۸/۸۰	۱۶۷۴/۳۳	۱۶۹۳/۷۱	۱۶۸۱/۰۴	۳۳/۰۱	شهریور
۴۳	۱۶۳۰/۶۹	۱۷۵۸/۸۹	۱۶۶۰/۰۸	۱۶۷۶/۰۴	۱۶۹۴/۲۹	۱۶۸۱/۸۸	۳۳/۱۶	مهر
۳۸	۱۶۳۰/۵۸	۱۷۵۷/۵۸	۱۶۶۰/۵۰	۱۶۷۴/۴۸	۱۶۹۱/۸۵	۱۶۸۲/۹۵	۳۹/۱۳	آبان
۴۴	۱۶۳۰/۸۷	۱۷۵۷/۷۷	۱۶۶۱/۸۰	۱۶۷۷/۲۵	۱۶۹۵/۱۹	۱۶۸۲/۷۶	۳۲/۷۴	آذر
۴۳	۱۶۳۰/۹۳	۱۷۵۷/۵۸	۱۶۶۱/۲۱	۱۶۷۷/۶۸	۱۶۹۵/۳۱	۱۶۸۲/۶۰	۳۲/۸۷	دی
۳۹	۱۶۳۱/۳۳	۱۷۵۷/۶۲	۱۶۶۴/۹۲	۱۶۸۱/۶۵	۱۶۹۴/۱۵	۱۶۸۷/۱۹	۳۷/۱۰	بهمن
۴۲	۱۶۳۱/۳۸	۱۷۵۷/۷۵	۱۶۶۳/۸۸	۱۶۷۸/۳۶	۱۶۹۵/۲۱	۱۶۸۲/۹۵	۳۲/۷۲	اسفند

در نرم‌افزار Matlab مشخص شده و ماتریسی ۱۲×۱۲ که سطرهای آن مقادیر برداشت‌شده در ماه‌های مختلف و ستون‌ها چاه‌های موجود در همسایگی هستند، ایجاد شد. جدول ۲ ماتریس تشکیل شده برای چاه ۱ را نشان می‌دهد. چاه‌های ۲-۶، ۸، ۱۰-۱۳، ۱۵ و ۲۰ چاه‌های موجود در همسایگی چاه ۱ هستند.

با توجه به اینکه چاه‌هایی که در روش PCA لحاظ می‌شوند، باید در همه ماه‌ها مقدار داشته باشند، برای چاه‌هایی که برداشت اطلاعات در ماه‌های خاصی انجام نشده بود با استفاده از افزونه xlstat در اکسل و روش NIPALS داده‌های جافتاده جدید تولید شدند. در پایان این مرحله، ۴۷ چاه نمونه‌برداری در ۱۲ ماه از سال دارای داده شدند. برای پیاده‌سازی مرحله دوم ابتدا ۱۲ همسایگی هر چاه

جدول ۲. ماتریس ایجادشده برای چاه ۱ بر اساس همسایگی اقلیدسی چاه‌ها (اعداد برحسب متر)

	۲۰	۱۵	۱۳	۱۲	۱۱	۱۰	۸	۶	۵	۴	۳	۲	
فروردین	۱۷۵۳/۹۹	۱۶۴۷/۵۳	۱۶۸۲/۶۲	۱۶۸۲/۰۰	۱۶۶۴/۷۱	۱۷۵۸/۱۶	۱۶۳۱/۳۳	۱۶۴۸/۲۷	۱۶۸۳/۲۲	۱۶۶۵/۱۷	۱۶۷۳/۴۷	۱۶۶۵/۲۵	
اردیبهشت	۱۷۵۳/۸۵	۱۶۴۷/۷۱	۱۶۹۵/۳۶	۱۶۸۲/۱۵	۱۶۶۲/۸۰	۱۷۶۰/۳۳	۱۶۳۱/۵۵	۱۶۴۸/۲۶	۱۶۸۳/۳۳	۱۶۶۵/۱۸	۱۶۷۳/۹۶	۱۶۶۵/۱۹	
خرداد	۱۷۵۳/۸۰	۱۶۴۷/۴۵	۱۶۹۵/۳۲	۱۶۸۰/۹۰	۱۶۶۲/۰۷	۱۷۶۲/۱۴	۱۶۳۱/۳۷	۱۶۴۸/۲۴	۱۶۷۹/۷۹	۱۶۶۵/۰۲	۱۶۷۳/۹۵	۱۶۶۵/۱۳	
تیر	۱۷۵۳/۶۴	۱۶۴۷/۱۴	۱۶۹۴/۰۸	۱۶۷۹/۳۸	۱۶۶۰/۸۴	۱۷۶۱/۲۹	۱۶۳۱/۰۴	۱۶۴۷/۲۶	۱۶۷۷/۷۴	۱۶۶۴/۸۷	۱۶۷۳/۵۸	۱۶۶۵/۱۱	
مرداد	۱۷۵۳/۶۳	۱۶۴۷/۰۱	۱۶۹۷/۹۶	۱۶۷۸/۶۹	۱۶۶۰/۸۱	۱۷۵۹/۷۸	۱۶۳۰/۸۳	۱۶۴۸/۲۴	۱۶۷۶/۴۱	۱۶۶۴/۶۹	۱۶۷۳/۴۱	۱۶۶۴/۸۳	
شهریور	۱۷۵۳/۸۲	۱۶۴۶/۸۸	۱۶۹۳/۷۱	۱۶۷۴/۳۳	۱۶۵۸/۸۰	۱۷۵۹/۲۰	۱۶۳۰/۷۰	۱۶۴۸/۲۲	۱۶۷۵/۸۰	۱۶۶۴/۵۳	۱۶۷۳/۲۱	۱۶۶۴/۷۸	
مهر	۱۷۵۳/۹۹	۱۶۴۶/۹۳	۱۶۹۴/۲۹	۱۶۷۶/۰۴	۱۶۶۰/۰۸	۱۷۵۸/۸۹	۱۶۳۰/۶۹	۱۶۴۸/۲۱	۱۶۷۷/۱۳	۱۶۶۴/۴۳	۱۶۷۳/۱۶	۱۶۶۴/۷۷	
آبان	۱۷۵۴/۰۹	۱۶۴۷/۲۱	۱۶۹۱/۸۵	۱۶۷۴/۴۸	۱۶۶۰/۵۰	۱۸۲۵/۵۸	۱۶۲۵/۵۸	۱۶۴۸/۴۴	۱۶۷۹/۱۵	۱۶۶۴/۳۷	۱۶۷۳/۱۵	۱۶۶۴/۷۶	
آذر	۱۷۵۴/۱۷	۱۶۴۷/۲۲	۱۶۹۵/۱۹	۱۶۷۷/۲۵	۱۶۶۱/۸۰	۱۷۵۷/۷۷	۱۶۳۰/۸۷	۱۶۴۸/۱۶	۱۶۸۰/۵۴	۱۶۶۴/۳۳	۱۶۷۳/۱۴	۱۶۶۴/۷۵	
دی	۱۷۵۳/۹۸	۱۶۴۷/۲۷	۱۶۹۵/۳۱	۱۶۷۷/۶۸	۱۶۶۱/۲۱	۱۷۵۷/۵۸	۱۶۳۰/۹۳	۱۶۴۸/۱۸	۱۶۸۱/۲۶	۱۶۶۴/۳۷	۱۶۷۳/۲۱	۱۶۶۴/۸۷	
بهمن	۱۷۵۴/۳۲	۱۶۴۷/۳۶	۱۶۹۴/۱۵	۱۶۸۱/۶۵	۱۶۶۴/۹۲	۱۸۲۴/۲۹	۱۶۳۱/۳۳	۱۶۴۸/۱۵	۱۶۸۱/۸۲	۱۶۶۴/۴۰	۱۶۷۳/۳۰	۱۶۶۴/۹۱	
اسفند	۱۷۵۳/۷۵	۱۶۴۷/۵۳	۱۶۹۵/۲۱	۱۶۷۸/۳۶	۱۶۶۳/۸۸	۱۷۵۷/۷۵	۱۶۳۱/۳۸	۱۶۴۸/۱۳	۱۶۸۲/۰۶	۱۶۶۴/۳۹	۱۶۷۳/۴۴	۱۶۶۴/۹۲	

جدول ۴ اهمیت نسبی هر چاه را نشان می‌دهد. هر چاه که اهمیت کمتری داشته باشد، رتبه کمتری دارد.

جدول ۳. ماتریس عاملی حاصل از اعمال روش PCA برای چاه ۱

شماره چاه	۱	۲	۳
۲	۰/۹۶		
۳	۰/۹۳		
۴	۰/۹۳		
۵	۰/۴۴	۰/۸۸	
۶	-۰/۲۷	۰/۴۲	۰/۸۳
۸	۰/۹۴		
۱۰	۰/۷۹	۰/۵۶	
۱۱	۰/۶۶	-۰/۶۵	
۱۲	۰/۷۵	-۰/۵۹	
۱۳	۰/۷۶		-۰/۴۶
۱۵	۰/۶۶	۰/۷۲	
۲۰	-۰/۴۲	۰/۶۶	

با اعمال آنالیز PCA روی تک تک ماتریس‌های ایجادشده در مرحله دوم، سهم متغیرها در مؤلفه‌ها محاسبه می‌شود. این مرحله در نرم‌افزار SPSS اجرا شد. نمونه‌ای از خروجی آن در جدول ۳ ارائه شده است، که در آن اعداد بزرگ‌تر یا کوچک‌تر از ۰/۳ به دلیل اینکه بدون تأثیر هستند، نوشته نشده‌اند. تعداد مؤلفه‌های اصلی در PCA برابر با تعداد چاه‌های استفاده‌شده در همسایگی هستند. مؤلفه اول PCA بیشترین واریانس را دارد و به ترتیب واریانس در مؤلفه‌های بعدی کاهش خواهد یافت. مقدار بیشینه سهم متغیرها در مؤلفه‌های اصلی^۱ در سه مؤلفه اول خود را نمایان می‌سازد. از این‌رو، سه ستون اول می‌توانند ملاک اصلی تعیین چاه‌های مؤثر بر شبکه باشند.

برای محاسبه اهمیت نسبی چاه‌ها، ابتدا تعداد دفعاتی که هر چاه ضریب همبستگی بیش از ۰/۹ با مؤلفه‌های PCA داشته باشد، شمارش شده و بر تعداد دفعاتی که هر چاه در همسایگی چاه‌های دیگر شمارش شده است، تقسیم می‌شود.

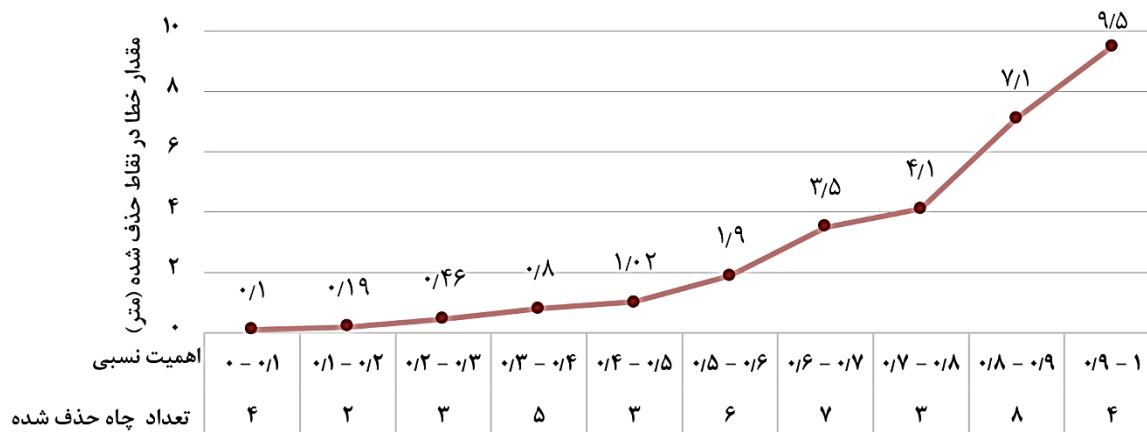
جدول ۴. اهمیت نسبی محاسبه شده برای هر نقطه بر اساس دفعات تکرار و دفعات مؤثر بودن

چاه	حضور در همسایگی	حضور به عنوان چاه مؤثر	اهمیت نسبی	چاه	حضور در همسایگی	حضور به عنوان چاه مؤثر	اهمیت نسبی
۱	۹	۳	۰/۳۳	۲۵	۹	۳	۰/۳۳
۲	۱۱	۷	۰/۵۴	۲۶	۱۳	۷	۰/۵۴
۳	۱۱	۲	۰/۱۱	۲۷	۱۸	۲	۰/۱۱
۴	۹	۰	۰	۲۸	۱۳	۰	۰
۵	۹	۷	۰/۱	۲۹	۷	۷	۰/۱
۶	۱۰	۷	۰/۳۵	۳۰	۲۰	۷	۰/۳۵
۷	۱۸	۱۲	۰/۵۷	۳۱	۲۱	۱۲	۰/۵۷
۸	۹	۱	۰/۰۶	۳۲	۱۶	۱	۰/۰۶
۹	۱۲	۵	۰/۶۳	۳۳	۸	۵	۰/۶۳
۱۰	۸	۸	۰/۷۳	۳۴	۱۱	۸	۰/۷۳
۱۱	۱۳	۱۲	۰/۶۳	۳۵	۱۹	۱۲	۰/۶۳
۱۲	۱۴	۶	۰/۵۵	۳۶	۱۱	۶	۰/۵۵
۱۳	۱۲	۹	۰/۹۰	۳۷	۱۰	۹	۰/۹۰
۱۴	۱۰	۵	۰/۳۶	۳۸	۱۴	۵	۰/۳۶
۱۵	۱۳	۸	۰/۸۹	۳۹	۹	۸	۰/۸۹
۱۶	۱۵	۳	۰/۲۳	۴۰	۱۳	۳	۰/۲۳
۱۷	۱۴	۱	۰/۰۹	۴۱	۱۱	۱	۰/۰۹
۱۸	۱۵	۶	۰/۸۶	۴۲	۷	۶	۰/۸۶
۱۹	۱۲	۱	۰/۱۱	۴۳	۹	۱	۰/۱۱
۲۰	۸	۷	۰/۸۸	۴۴	۸	۷	۰/۸۸
۲۱	۱۴	۵	۰/۴۲	۴۵	۱۲	۵	۰/۴۲
۲۲	۱۱	۵	۰/۶۳	۴۶	۸	۵	۰/۶۳
۲۳	۱۸	۳	۰/۵۰	۴۷	۶	۳	۰/۵۰
۲۴	۱۶	۹	۰/۵۶				

در مرحله بعدی مشخص می‌شود که مقادیر چاه‌های کم‌اهمیت که در اولویت حذف هستند، توسط ایستگاه همسایه با چه دقتی قابل اندازه‌گیری است و حذف این ایستگاه‌ها چه مقدار خطا در پهنه‌بندی داده‌ها ایجاد می‌کند. شکل ۴ متوسط مقدار خطای برآورد شده در صورت حذف ایستگاه‌های نمونه‌برداری با اولویت‌های کم در کل ۱۲ ماه را نشان می‌دهد. برای ایجاد این شکل تعداد چاه‌ها با اهمیت‌های نسبی مختلف دسته‌بندی شده و به ازای حذف چاه‌ها هر دسته مقدار خطا محاسبه شده است.

همان‌گونه از جدول ۴ مشخص است، چاه‌های ۲۰ و ۷ دارای بیشترین ارزش نسبی (مقدار ۱) بوده و یک چاه دارای اهمیت نسبی صفر است. چاه ۲۸ در آنالیز PCA در ماتریسی عاملی هیچ چاهی مقدار بیش از ۰/۹ نداشته‌اند. بنابراین، نشان می‌دهد که در محاسبات و اطلاعات سایر چاه‌ها کمترین نقش را داشته است. دلیل این امر می‌تواند موقعیت مکانی چاه ۲۸ باشد، چرا که در بخش مرکزی آبخوان تراکم نقاط بیشتر از سایر بخش‌ها است.

تا این مرحله اولویت نسبی چاه‌ها با توجه به اطلاعاتی که طی یک سال به شبکه پایش وارد می‌کردند، بررسی شدند.



شکل ۴. متوسط تغییرات ایجادشده در صورت حذف چاه با اولویت

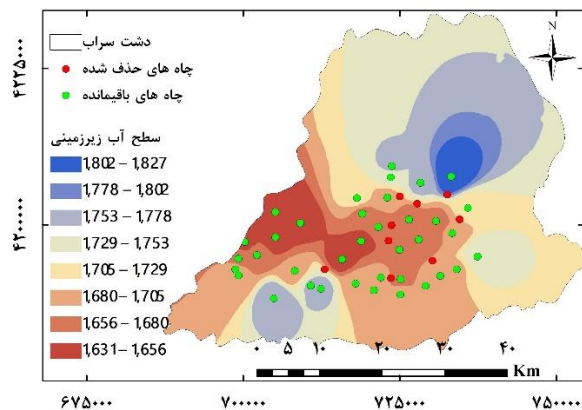
۰/۳ مقدار RMSE تغییر چندانی نمی‌کند، یعنی اگر حداکثر خطای RMSE ۰/۵ متر قابل قبول باشد، با حذف چاه‌های کم‌اهمیت خطای چندانی در تخمین سطح آب زیرزمینی آبخوان رخ نمی‌دهد. ولی اگر چاه‌ها با اولویت نسبی بیشتر از ۰/۳ حذف شوند، مقدار خطا (RMSE) زیاد می‌شود. بنابراین، در این تحقیق بر اساس نظرات کارشناسی و معادل‌سازی حدود آستانه خطای تحقیقات پیشین مقدار خطای ۰/۵ متر در چاه‌های حذف‌شده قابل قبول لحاظ شد. در این آستانه از ۴۷ حلقه چاه، ۳۸ حلقه به‌عنوان چاه مؤثر شناسایی شده و بقیه چاه‌ها به‌عنوان چاه‌های کم‌اهمیت شناخته می‌شوند. به این ترتیب، با ۳۸ حلقه چاه در منطقه می‌توان در وقت و هزینه صرفه‌جویی کرد و دقت زیادی در اندازه‌گیری و قرائت چاه‌های مهم برای افزایش دقت مدل‌سازی به عمل آورد. بررسی موقعیت مکانی چاه‌های حذف‌شده نشان می‌دهد بیشتر این چاه‌ها در مناطقی قرار دارند که در مجاورت آنها تراکم چاه‌ها بیشتر است. زیاد بودن تراکم در اطراف چاه حذف‌شده کمک می‌کند که سطح آب زیرزمینی در آن نقطه با توجه به ضریب همبستگی مکانی موجود بین چاه‌ها، با استفاده از مقادیر چاه‌های دیگر محاسبه شود. تراکم کم نقاط در بخش جنوب غربی مانع از حذف چاه‌های نمونه‌برداری شده، به گونه‌ای که از این بخش هیچ چاهی حذف نشده است.

شکل ۵ خروجی روش درون‌یابی کریجینگ در ماه فروردین با استفاده از کل نقاط و همچنین، با فرض حذف ۹ چاه با اولویت کمتر از ۰/۳ را نشان می‌دهد. مقدار خطای

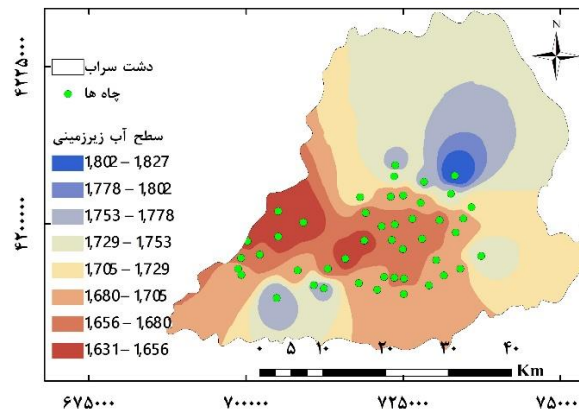
بر اساس شکل ۴ با حذف چاه‌های با اولویت کمتر از ۰/۱ (بازه ۰/۱) که ۴ چاه است، ۰/۱ متر خطا در چاه‌های حذف‌شده ایجاد می‌شود و به این ترتیب، با افزایش تعداد چاه‌های حذف‌شده، این خطا افزایش پیدا می‌کند. در تحقیقات پیشین چاه‌ها با اهمیت نسبی متفاوتی حذف شده‌اند و مقادیر مختلفی به‌عنوان بیشترین خطای قابل قبول در سطوح جدید ایجادشده معرفی شده است. کماسی و گودرزی (۱۴۰۰) دقت RMSE ای برابر با ۰/۶۰۵ متر برای پهنه‌بندی نقاط بدون آمار لحاظ کردند [۱۰]. خاشعی سیوکی و همکاران (۱۴۰۰) حد آستانه ۱۲ درصد خطا در ضریب تغییرات را ملاک قرار دادند و چاه‌های با اولویت نسبی کمتر از ۰/۵ را حذف کردند [۹]. صیادی شهرکی و همکاران (۱۳۹۹) چاه‌ها با اولویت نسبی کمتر از ۰/۸ را حذف کردند [۱۱]. نوری قیداری (۱۳۹۲) چاه‌های با اولویت نسبی کمتر از ۰/۵ که موجب خطای ۱۳ درصد می‌شد را حذف کرد [۱۸]. تعداد چاه حذف‌شده نیز در بهینه‌سازی شبکه پایس سطح آب زیرزمینی در تحقیقات مختلف متفاوت بوده است. مثلاً در آبخوان دشت سیلاخور ۶۰ درصد از چاه‌ها قابل حذف بودند [۱۰]. در آبخوان دشت شهرکرد، ۱۷ درصد [۵]؛ در آبخوان دشت دزفول-اندیمشک، ۲۰ درصد [۲۴] و در آبخوان دشت دهگلان، ۱۹ درصد [۷] چاه‌ها قابل حذف معرفی شدند. با توجه به تحقیقات پیشین، بیشترین خطای قابل قبول و تعداد چاه‌های حذف‌شده بر اساس نظرات کارشناسی، تراکم اولیه شبکه، دقت مورد نیاز در منطقه و روش استفاده‌شده تعیین می‌شود. از طرفی، با توجه به شکل ۴ در اولویت نسبی ۰-

در نقاط حذف‌شده ایجاد می‌شود. بررسی بصری دو سطح نیز نشان می‌دهد که تغییرات چندانی در مقادیر دو سطح وجود نداشته است.

ایجادشده در ماه فروردین برای ایستگاه‌های حذف‌شده $RMSE=0/41m$ است که نشان می‌دهد با حذف ۹ چاه از مجموعه چاه‌های موجود، در ماه فروردین ۴۱ سانتی‌متر خطا



سطح ایجاد شده بعد از حذف چاه‌های با اولویت نسبی کمتر از ۰/۳



سطح ایجادشده با تمام نقاط

شکل ۵. سطوح پیوسته ایجادشده در دشت سراب

استفاده از مزایای دو روش زمین‌آماری و PCA و همچنین کاهش ضعف‌های این دو روش، رویکردی تلفیقی ارائه شده و در تخمین سطح آب زیرزمینی آبخوان دشت سراب، قابلیت آن بررسی شد. در رویکرد پیشنهادی اطلاعات مکانی و زمانی چاه‌ها به صورت مؤثر بر فرایند تعیین چاه‌های تأثیرگذار استفاده می‌شود. به این منظور، ابتدا اهمیت نسبی چاه‌ها در روش PCA با لحاظ اطلاعات زمانی چاه‌های موجود در همسایگی محاسبه شده و سپس، با استفاده از ضریب همبستگی مکانی در روش زمین‌آماری کریجینگ مقدار خطای ایجادشده در شبکه محاسبه می‌شود.

تحلیل اکتشافی داده‌های ۱۲ ماه از سال ۱۳۹۷ در دشت سراب، دو چاه با مقادیر پرت را مشخص کرد که موجب شد از ادامه محاسبات خارج شوند. در روش PCA استفاده از اطلاعات چاه‌هایی امکان‌پذیر است که به صورت پیوسته برداشت اطلاعات در آنها انجام شده باشد. بنابراین، برای هر ماه داده‌های جاقفاده بازتولید شدند. نتایج اولیه از پیاده‌سازی روش PCA اهمیت نسبی چاه‌ها را مشخص کرد که بر اساس آن، دو چاه دارای بیشترین اهمیت نسبی و یک چاه بدون اهمیت محاسبه شدند. در ادامه، برای تعیین اینکه حذف چه تعداد از چاه‌ها مجاز است، از خطای پهنه‌بندی روش کریجینگ در چاه‌های با اهمیت نسبی کم با کمیت $RMSE$ استفاده شد. در این مرحله بیشترین خطای قابل قبول در

بر اساس نقاط حذف‌شده مشاهده می‌شود که بیشترین نقاط حذف‌شده در بخش مرکزی دشت قرار دارد، منطقه‌ای که فاصله مکانی بین داده‌ها کمتر است. همان‌گونه که در شکل ۵ مشخص است، توزیع مکانی نشان‌داده‌شده توانسته است برآورد مطلوبی را از سطح آب زیرزمینی در دشت سراب داشته باشد که موجب حداقل کاهش دقت در مدل‌سازی در نقاط بدون آمار شود. از طرف دیگر، شبکه جدید می‌تواند زمینه مناسبی را برای مدیریت منابع آب زیرزمینی، مدل‌سازی مکانی-زمانی آبخوان و به دنبال آن، بررسی و پیش‌بینی جریان آب زیرزمینی را فراهم آورد.

نتیجه‌گیری

در طراحی شبکه‌های پایش، انتخاب موقعیت مکانی و زمانی ایستگاه‌های با اهمیت به گونه‌ای که اطلاعات ضروری و نه مازاد را با کمترین هزینه در اختیار قرار دهند، همواره مورد توجه بوده است. شبکه بهینه‌شده نسبت به حالت اولیه، شبکه‌ای ناقص است که بهترین عملکرد را در بین شبکه‌های ناقص دارد. در بین روش‌های استفاده‌شده برای بهینه‌سازی شبکه پایش روش خاصی که برتری آن بر دیگر روش‌ها اثبات شده باشد، وجود ندارد و با توجه به موارد مختلفی از جمله پدیده بررسی‌شده، گستره منطقه و حجم داده‌های در دسترس روش مناسب انتخاب می‌شود. در این تحقیق برای

- [4]. Aadil N, Gallardo A, Ahmed S. Optimization of a Groundwater Monitoring Network for a Sustainable Development of the Maheshwaram Catchment, India. *Sustainability*. 2011;3.
- [5]. Abdollahi Mansourkhani M, Mohammadzade H, Amini M, Azizi F. Assessment of Groundwater Quality Spatial Distribution and Appointment Optimize Network of Shahrkord Plain Aquifer Using Geostatistical Methods. *Watershed Management Research Journal*. 2019;32(2):60-78.
- [6]. Taheri Zangi S, Vaezihir A. Vulnerability of Shazand Plain Subsidence Caused by Groundwater Level Reduction Using Weighting Model and Its Validation Analysis Using Radar Interferometry. *Iranian journal of Ecohydrology*. 2020;7(1):183-94.
- [7]. Hosseini M, Kerachian R. A data fusion-based methodology for optimal redesign of groundwater monitoring networks. *Journal of Hydrology*. 2017;552:267-82.
- [8]. Lashkaripour G, Rostami Barani H, Kohandel A, Tarshizi H. Groundwater level drop and land subsidence in Kashmar plain. 10th Iranian Geological Society Conference; Tehran 2006.
- [9]. Khashei A, Shahidi A, Rahnama S. Comparison of Birjand Plain Aquifer Chromium Monitoring Network Using Principal Component Analysis (PCA) and Entropy Theory. *Environment and Water Engineering*. 2021;7(2):220-31.
- [10]. komasi m, goudarzi h. Multi-Objective Optimization Groundwater Network Using Genetic Algorithm (NSGA-II) and Empirical Bayesian Kriging (EBK) Method (Case Study: Silakhor plain). *Irrigation and Water Engineering*. 2021;11(3):204-20.
- [11]. Shahidi A, Khashei Siouki A, Ramezani Y, Nazeri Tehrani M. Design of rain gauge monitoring network using irregularity theory (Case study: Urmia Lake Basin). *Irrigation and Drainage of Iran*. 2019;13(2):296-308.
- [12]. Vu MT, Jardani A, Massei N, Fournier M. Reconstruction of missing groundwater level data by using Long Short-Term Memory (LSTM) deep neural network. *Journal of Hydrology*. 2020;597:125776.
- [13]. Khodaverdi M, Hashemi sR, Khashei-Siuki A, Pourreza- Bilondi M. Optimal Design of Groundwater-Quality Sampling Networks with MOPSO-GS (Case Study: Neyshabour Plain). *Water and Irrigation Management*. 2020;9(2):199-210.

چاه‌های حذف‌شده و بیشترین تعداد چاه‌های حذف‌شده با توجه به تحقیقات پیشین و نظرات کارشناسی مشخص شد. بر اساس حد آستانه ۰/۵ متر خطا در چاه‌های حذف‌شده، از ۴۷ حلقه چاه سفره اول دشت سراب، ۳۸ چاه (۸۱ درصد) به‌عنوان چاه‌های مؤثر شناسایی شده و بقیه چاه‌ها به‌عنوان چاه‌های کم‌اهمیت شناخته شدند. بررسی چاه‌های حذف‌شده نشان می‌دهد که بیشتر در مرکز دشت سراب قرار دارند و از حاشیه آبخوان چاهی حذف نشده است. رویکرد پیشنهادی مزایای دو روش PCA و زمین‌آماری را دارد و با تلفیق مفاهیم روش‌ها، ضعف‌ها هر دو روش را تا حدی کاهش می‌دهد. روش پیشنهادی با توجه به ماهیت PCA که در آن باید اطلاعات چاه‌ها به صورت پیوسته برداشت شده باشند برای دشت‌هایی که برداشت اطلاعات به صورت مرتب اندازه‌گیری شده مفید است؛ اما این روش برای مناطقی که داده‌های جافتاده در چاه‌ها زیاد است، توصیه نمی‌شود. برای تحقیقات آتی پیشنهاد می‌شود تلفیق سایر روش‌های بهینه‌سازی شبکه پایش مانند آنتروپی و زمین‌آمار بررسی شده و خروجی آن با روش پیشنهادی این تحقیق مقایسه شود. پیشنهاد می‌شود بعد از تعیین اولویت نسبی چاه‌ها از روش‌های بهینه‌سازی مانند ژنتیک برای بهینه‌سازی ایستگاه‌های پایش استفاده شود. همچنین، توصیه می‌شود داده‌ها در سری زمانی طولانی‌تر مانند ۱۰ سال مد نظر قرار گیرد و روش ارائه‌شده در پدیده‌های مختلفی دیگری مانند ایستگاه‌های پایش پارامترهای هواشناسی بررسی شود.

منابع

- [1]. Sayadi shahraki A, naseri a, boromandnasab s, soltani a. Designing a network for monitoring groundwater level using the Principal Component Analysis technique. 2020;13(44):29-37.
- [2]. Destandau F, Zaiter Y. Spatio-temporal design for a water quality monitoring network maximizing the economic value of information to optimize the detection of accidental pollution. *Water Resources and Economics*. 2020;32:100156.
- [3]. Ou C-P, St-Hilaire A, Ouarda T, Conly FM, Armstrong N, Khalil B, et al. Coupling geostatistical approaches with PCA and fuzzy optimal model (FOM) for the integrated assessment of sampling locations of water quality monitoring networks (WQMN). *Journal of environmental monitoring : JEM*. 2012;14.

- [23]. Silva M, Santos A, Santos R, Figueiredo E, Sales C, Costa JCWA. Deep principal component analysis: An enhanced approach for structural damage identification. *Structural Health Monitoring*. 2018;18(5-6):1444-63.
- [24]. Raeisi A, Ghafouri H-R, Moslemzadeh M. Minimization of Groundwater Observation Wells Using Geostatistics and Optimization Technique (Case study: Dezfoul-Andimeshk plain). *Journal of Water and Soil Conservation*. 2018;25(3):79-96.
- [25]. Hooshangi N. Determination of valuable piezometric wells in groundwater level prediction by considering spatiotemporal information. *GEO*. 2020;13(49):37-49.
- [26]. Hooshangi N, Alesheikh A, Nadiri A. Optimization of Piezometers Number for Groundwater Level Prediction Using PCA and Geostatistical Methods. *Water and Soil Science*. 2016;25(4/2):53-66.
- [27]. Jahanbakhsh Asl S, Sari Sarraf B, Khorshid Doost AM, Rostamzadeh H. Evaluation of vegetation changes in Sarab plain and analysis of drought and wet seasons *Geography* 2009;23(7):118-34.
- [28]. karami F, Rostamzadeh H. Investigation of effective factors in salinization of Sarab plain lands. *Iranian Journal of Natural Resources*. 2008;4(3).
- [29]. Kaffash Charandabi N. Forecasting Air Pollution based on Monitoring Station with using Kalman Filter. *New Approaches in Civil Engineering*. 2019;3(3):46-60.
- [30]. Preda C, Saporta G, Mbarek M. The NIPALS algorithm for missing functional data. *Revue Roumaine de Mathématiques Pures et Appliquées*. 2010;55.
- [31]. Hooshangi N, Alesheikh AA. Evaluation OF ANN, ANFIS and FUZZY Systems in estimation of solar radiation in Iran. *Journal of gematics science and technology*. 2015;4(3):187-200.
- [14]. Wang C, Zhao L, Sun W, Xue J, Xie Y. Identifying redundant monitoring stations in an air quality monitoring network. *Atmospheric Environment*. 2018;190:256-68.
- [15]. Seifipour K, Mirabbasi R, Mirzaei M. Application of Entropy Theory in Assessing Groundwater Quality Monitoring Network of Sefiddasht. *Hydrogeology*. 2020;4(2):63-73.
- [16]. Sottani A, Meggiorin M, Ribeiro L, Rinaldo A. Comparison of two methods for optimizing existing groundwater monitoring networks: application to the Bacchiglione Basin, Italy2020.
- [17]. Galán-Madruga D, García-Camero JP. An optimized approach for estimating benzene in ambient air within an air quality monitoring network. *Journal of Environmental Sciences*. 2022;111:164-74.
- [18]. Noori gheidari Mh. Determination of Effective Wells to Monitor the Ground Water Level Using the Principal Components Analysis. *JSTNAR*. 2013;17(64):149-59.
- [19]. Camacho J, Pérez-Villegas A, García-Teodoro P, Maciá-Fernández G. PCA-based multivariate statistical network monitoring for anomaly detection. *Computers & Security*. 2016;59:118-37.
- [20]. Teng SY, How BS, Leong WD, Teoh JH, Siang Cheah AC, Motavasel Z, et al. Principal component analysis-aided statistical process optimisation (PASPO) for process improvement in industrial refineries. *Journal of Cleaner Production*. 2019;225:359-75.
- [21]. Ghadban N, Honeine P, Francis C, Mourad-Chehade F, Farah J, editors. Strategies for principal component analysis in wireless sensor networks. 2014 IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM); 2014 22-25 June 2014.
- [22]. Babaei Hesar S, Hamdami Q, Ghasemieh H. Identify the Effective Wells in Determination of Groundwater Depth in Urmia Plain Using Principle Component Analysis. *Water and Soil*. 2017;31(1):40-50.