



مدیریت آب و آبیاری

دوره ۱۲ ■ شماره ۲ ■ تابستان ۱۴۰۱

صفحه‌های ۲۶۲-۲۴۵

DOI: 10.22059/jwim.2022.339038.967

مقاله پژوهشی:

بازنگری شبکه پایش کیفی آب زیرزمینی با ترکیب روش‌های خوشه‌بندی و زمین‌آمار در محدوده مطالعاتی تهران-کرج

یاسمن ساده^۱، علی مریدی^{۲*}، سید سعید موسوی ندوشنی^۳

۱. دانشجوی کارشناسی ارشد، گروه مهندسی منابع آب، دانشکده عمران، آب و محیط زیست، دانشگاه شهید بهشتی، تهران، ایران.

۲. استادیار، گروه مهندسی محیط زیست، دانشکده عمران، آب و محیط زیست، دانشگاه شهید بهشتی، تهران، ایران.

۳. دانشیار، گروه مهندسی منابع آب، دانشکده عمران، آب و محیط زیست، دانشگاه شهید بهشتی، تهران، ایران.

تاریخ دریافت مقاله: ۱۴۰۰/۱۱/۲۴

تاریخ پذیرش مقاله: ۱۴۰۰/۱۲/۷

چکیده

حراست از کمیت و کیفیت منابع آب، همواره و در تمام جوامع بشری دارای اهمیت فراوانی بوده است و به منظور حفظ کیفیت این منابع، پایش‌ها و اقدامات اصلاحی متعددی در اکثر کشورهای جهان صورت گرفته است. در این راستا پایش کیفیت آب به‌عنوان یکی از ابزارهای ضروری و به‌عنوان یک فعالیت یکپارچه برای ارزیابی عوامل فیزیکی، شیمیایی و بیولوژیکی آب که با سلامت انسان و موجودات زنده رابطه دارند، مطرح می‌باشد. در پژوهش حاضر برای بازنگری و ارائه شبکه پایش کیفی منابع آب زیرزمینی مدلی از ترکیب روش‌های زمین‌آمار کریجینگ، خوشه‌بندی و تئوری آنتروپی ارائه شده است. این مدل تحت دو رویکرد اول و دوم در محدوده مطالعاتی تهران-کرج ارائه شده است. رویکرد اول بدون استفاده از روش خوشه‌بندی فقط با استفاده از تئوری آنتروپی و روش زمین‌آمار کریجینگ به‌عنوان تخمین‌گر، به بازنگری شبکه پایش موجود پرداخته است. رویکرد دوم با استفاده از روش خوشه‌بندی k-means، تئوری آنتروپی و روش زمین‌آمار کریجینگ به‌عنوان تخمین‌گر، تأثیر ترکیب این سه روش را بر بازنگری شبکه پایش موجود بررسی کرده است و سپس نتایج رویکرد اول و دوم مقایسه شده‌اند. شبکه پایش نهایی پیشنهاد شده با تعداد ۴۴ حلقه چاه، متوسط درصد خطای تخمین ۱۹ و با کاهش هزینه ۳۴ درصدی نسبت به هزینه شبکه پایش فعلی ارائه شده است. همچنین با استفاده از خوشه‌بندی متوسط درصد خطای تخمین نسبت به حالت بدون خوشه‌بندی ۲۰ درصد کاهش یافته است.

کلیدواژه‌ها: آب زیرزمینی، پایش کیفی، تئوری آنتروپی، کریجینگ، K-means.

Review of groundwater quality monitoring network by combining clustering and geostatistical methods in Tehran-Karaj study area

Yasaman Sadeh¹, Ali Moridi^{2*}, Seyed Saied Mousavi Nadoushani³

1. M.Sc. Student, Department of Water Resources Engineering, Faculty of Civil, Water and Environmental Engineering, Shahid Beheshti University, Tehran, Iran.

2. Assistant Professor, Department of Environmental Engineering, Faculty of Civil, Water and Environmental Engineering, Shahid Beheshti University, Tehran, Iran.

3. Associate Professor, Department of Water Resources Engineering, Faculty of Civil, Water and Environmental Engineering, Shahid Beheshti University, Tehran, Iran.

Received: February 13, 2022

Accepted: February 26, 2022

Abstract

Protecting the quantity and quality of water resources has always been of great importance in all human societies, and in order to maintain the quality of these resources, numerous monitoring and remedial measures have been taken in most countries of the world. In this regard, water quality monitoring is considered as one of the essential tools and as an integrated activity to evaluate the physical, chemical and biological factors of water that are related to human health and living organisms. In the present study, a model of combining geostatistical methods (Kriging), clustering and entropy theory has been proposed to review and present the groundwater quality monitoring network. This model is presented under the first and second approaches in Tehran-Karaj study area. The first approach, without using the clustering method, reviews the existing monitoring network without using only entropy theory and Kriging geostatistical method as an estimator. The second approach uses the k-means clustering method, entropy theory and Kriging geostatistical method as an estimator to investigate the effect of combining these three methods on the review of the existing monitoring network and then the results of the first and second approaches are compared. The proposed final monitoring network with 44 wells has an average forecast error rate of 19 and a cost reduction of 34 percent compared to the cost of the current monitoring network. Also, using clustering, the average percentage of estimation error has been reduced by 20 percent compared to the case without clustering.

Keywords: Entropy theory, Groundwater, K-means, Kriging, Quality monitoring.

مقدمه

در کشور ایران، منابع آب زیرزمینی به‌عنوان مهم‌ترین منابع تأمین آب موردنیاز برای بخش‌های کشاورزی، شرب و صنعت از اهمیت زیادی برخوردار است. کشاورزی با اختصاص سهم ۹۳ درصدی و برداشت بیش از ۸۰ درصد آن از منابع آب زیرزمینی نقش عمده‌ای در تغییرات کمی و کیفی آبخوان‌ها دارد (Ahmadi & Sedghamiz, 2007). مطالعاتی که در سال‌های اخیر انجام شده است، نشان‌گر تغییرات محسوس در بسیاری از مناطق کره زمین می‌باشد، که بر کمیت و کیفیت آب تأثیرگذارند. به‌منظور مدیریت کیفیت آب در مخازن سدها، تالاب‌ها، حوضچه‌ها و رودخانه‌ها شناخت پدیده‌هایی که در آن‌ها رخ می‌دهند ضروری است. آگاهی از وضعیت کیفیت آب‌های زیرزمینی، تغییرات و کنترل آن یکی از نیازهای مدیریتی بوده و بسیار دارای اهمیت می‌باشد. آب‌های زیرزمینی منبع مهمی از آب هستند، به‌ویژه در مناطق خشک و نیمه‌خشک که آب‌های سطحی کمیاب هستند. در نتیجه آلودگی آب‌های زیرزمینی در این مناطق یک نگرانی بزرگ است، به‌ویژه که کنترل و حذف آلودگی در این منابع نه‌تنها پرهزینه بلکه گاهی غیرممکن است (Rezaei et al., 2013). افزایش تقاضا برای آب، استانداردهای بالاتر زندگی، کاهش منابع با کیفیت قابل قبول و آلودگی بیش از حد آب به‌دلیل گسترش شهرها، کشاورزی و صنعت باعث ایجاد مخمصه‌های شدید زیست‌محیطی، اجتماعی، اقتصادی و سیاسی شده است (Karamouz et al., 2020). بنابراین، لازم است مدیریت کیفیت منابع آب در سطح حوضه آبریز از مبدأ تا مصرف، با انجام اقدام‌های سازه‌ای و غیرسازه‌ای به‌منظور حفاظت، بهبود و کنترل کیفیت منابع آب برای تأمین مطلوب نیازهای کلیه بهره‌برداران صورت پذیرد که درواقع پایش کیفیت آب یکی از این اقدام‌هاست. پایش

کیفیت آب فرایند برنامه‌ریزی‌شده نمونه‌برداری، اندازه‌گیری و ثبت یا درج ویژگی‌های مختلف آب است که به‌منظور تجزیه و تحلیل اطلاعات و بررسی روند تغییرات، ارزیابی تناسب و تطابق برنامه با هدف یا اهداف کاربری‌های تعریف‌شده برای آب و بررسی پای‌بندی و رعایت ضوابط توسط مصرف‌کنندگان آب صورت می‌گیرد.

Alfonso et al. (2010) به‌منظور یافتن راحت‌ترین مجموعه مکان‌ها برای قرارگیری ایستگاه‌های پایش، یک مسأله‌مسئله بهینه‌سازی چندهدفه تحت در نظر گرفتن دو ملاحظه در نظر گرفتن هزینه قرارگیری ایستگاه‌های پایش جدید و در نظر گرفتن هزینه ایستگاه‌های پایش نزدیک به سازه‌های هیدرولیکی، مطرح کرده‌اند. در هر دو حالت آنتروپی بیشینه شده و همبستگی کل نیز کمینه شده است. روش پیشنهادی در مطالعه موردی منطقه دلفند هلند به‌کار گرفته شده است و نتایج نشان داده است که استفاده از همبستگی کل روش مؤثری برای اندازه‌گیری استقلال چندمتغیره است و باید با آنتروپی مشترک ترکیب شود تا نتایجی که به‌دست می‌آید، بخش قابل‌توجهی از کل محتوای سیستم را پوشش دهد. بیشینه‌کردن آنتروپی مشترک نتایجی را مبنی بر پوشش ۸۲ تا ۸۵ درصد از کل محتوای اطلاعاتی را داده است. Daughney et al. (2012) از روشی برای مشخص کردن شبکه پایش آب زیرزمینی استفاده کرده‌اند. از روش تجزیه و تحلیل خوشه‌ای سلسله مراتبی (HCA) به‌عنوان ابزاری برای ارزیابی شبکه پایش آب زیرزمینی در منطقه‌ای از نیوزلند استفاده کرده‌اند. روش HCA شبکه‌های پایش نیوزلند را از نظر تعداد خوشه‌های کیفیت آب شناسایی‌شده در هر شبکه، متغیر هیدروشیمیایی مرکز هر دسته و محدوده غلظت هر ماده شیمیایی در هر خوشه کیفی آب مقایسه می‌کند. این روش می‌تواند برای ارزیابی هر زیرمجموعه سایت‌های پایش که

ثابت‌ها و متغیرهای مدل) که در طراحی شبکه پایش تأثیرگذار است براساس اطلاعات موجود و مشاوره با کارشناسان مشخص شده‌اند. برای تعیین وزن عوامل از فرایند سلسله مراتبی تحلیلی^۲ (AHP) استفاده شده است. اولین قدم در AHP انتخاب معیارهای مربوط به توزیع مکانی نقاط پایش در شبکه است. دلیل انتخاب این معیارها امکان ارزیابی میزان استفاده بیش از حد از سفره آبخوان و تأثیرات آن است. وزن‌دهی به عوامل نیز باید با نظر کارشناسان در آب زیرزمینی و GIS انجام گیرد. سپس روش ترکیبی خطی وزنی^۳ (WLC) برای تعیین اولویت‌بندی مناطق برای پایش استفاده شده است. همانند سایر روش‌ها که دارای محدودیت هستند، AHP نیز از این قاعده مستثنی نیست چون تحلیل AHP معمولاً مبتنی بر فرایندهای طولانی و تکراری است که به زمان زیادی برای رسیدن به یک نتیجه در مورد معیارهای موردبررسی و وزن آن‌ها نیاز دارد. این روش می‌تواند در آبخوان‌هایی که ویژگی‌های متنوعی دارند یا بیش‌ترین محدودیت آن‌ها عدم وجود داده‌های قابل‌اعتماد است، استفاده شود. یکی از مزایای این روش این است که تصمیم‌گیری مسائل پیچیده را راحت می‌کند. به‌منظور پیداکردن نقشه اولویت‌بندی مناطق برای پایش، نقشه به‌دست‌آمده دوباره طبقه‌بندی^۴ می‌شود. این طبقه‌بندی پنج کلاس خیلی کم تا خیلی زیاد است، که ناحیه خیلی زیاد در منطقه شهری قرار گرفته که اولویت بیش‌تری برای پایش دارد و داده‌های بیش‌تری در مورد پویایی آبخوان در اختیار قرار داده است. نقشه نهایی از نظر مکانی با رفتار معیارهای مورداستفاده در پژوهش، سازگار بوده است. *Du et al.* (2017) همبستگی بین متغیرهای کیفی آب‌های ساحلی را تحلیل کرده‌اند. ارزیابی کیفیت آب برای ارزیابی پیرآبی دریایی، تخمین رشد جلبک‌های مضر و حفظت از محیط زیست ضروری است و روشی برای ارزیابی کیفیت آب

از یک شبکه پایش بزرگ‌تر گرفته شده است، استفاده شود. *Alfonso et al.* (2013) دو روش برای طراحی شبکه‌های پایش دبی رودخانه‌ها با استفاده از مفاهیم تئوری اطلاعات ارائه کرده‌اند. روش اول بهینه‌سازی مقادیر تئوری اطلاعات (با استفاده از الگوریتم ژنتیک چندهدفه) و روش دوم روش جدیدی مبتنی بر رتبه‌بندی مقادیر تئوری اطلاعات با ترکیب‌های مختلف پایش در نظر گرفته شده است. این روش‌ها برای رودخانه مگدلینا در کلمبیا به‌کار گرفته شده است و شبکه پایش موجود نیز ارزیابی شده است و نشان داده شده است که روش رتبه‌بندی روشی امیدوارکننده برای یافتن نقاط حداکثر جبهه پارتو است که طی بهینه‌سازی چندهدفه به‌دست آمده است. *Chang & Lin* (2014) یک تحلیل چندمعیاره برای ارزیابی تناسب طراحی یک شبکه پایش کیفی آب ارائه داده‌اند. برای کنترل کیفیت آب یک حوضه آبریز، یک سیستم پایش کیفی آب لازم است و عوامل متعددی می‌توانند بر طراحی شبکه پایش کیفی مؤثر باشند. برای این تحلیل هفت معیار در نظر گرفته شده است که نرمال شده‌اند و به هرکدام یک وزن اختصاص داده می‌شود. روش وزن‌دهی برای نمره‌دهی به زیرحوضه‌ها استفاده می‌شود و در زیرحوضه‌ای که چگالی بیش‌تری داشته باشد، نیاز به پایش کیفی بیش‌تر است. سپس از روش فازی (Fuzzy) برای اولویت‌بندی نیاز به چگالی بیش‌تر برای ایستگاه‌های پایش کیفی آب استفاده می‌شود. با این روش اولویت‌بندی مناطق با چهار گروه مشخص می‌شود. نتایج نشان می‌دهد که بعضی از زیرحوضه‌ها پایش کیفی بیش‌تری نیاز دارند. *Esquivel et al.* (2015) یک تحلیل چندمعیاره برای تحلیل و مدل‌کردن معیارهای اصلی تأثیرگذار بر طراحی بهینه شبکه پایش سطح آب زیرزمینی ارائه کرده‌اند. تحلیل چندهدفه با استفاده از GIS انجام گرفته است. تعاریف و اهمیت معیارها (یعنی انتخاب

آب یک خلیج در ایالات متحده آمریکا معرفی کرده‌اند. نیم‌وارپانس و آنتروپی انتقال اطلاعات گسسته برای محاسبه محدوده بهینه فاصله پایش اعمال شده‌اند. این مطالعه محدوده بهینه ۲۸ تا ۸۲ و ۳۷ تا ۵۰ کیلومتری را برای متغیرهای مطالعه شده نشان داده است. در صورت در نظر گرفتن فاصله پایش در فاصله‌های ذکر شده، اطلاعات مفیدی از شبکه پایش به دست می‌آید. Alizadeh *et al.* (2018) یک روش برای طراحی شبکه پایش کیفی بهینه براساس تئوری آنتروپی ارائه کرده‌اند. ابتدا مواد جامد محلول (TDS)^{۱۰} توسط روش‌های شبکه‌های عصبی مصنوعی (ANN)^{۱۱} و نزدیک‌ترین همسایگی (KNN)^{۱۲} تخمین زده می‌شود سپس با استفاده از الگوریتم بهینه‌سازی (Differential Evolution (DE)) با هدف ماکزیم‌سازی اطلاعات خالص که از ایستگاه‌ها به دست می‌آید، تعداد و محل ایستگاه‌های پایش نیز براساس چاه‌های موجود (فعال) به دست می‌آید. توزیع مکانی چاه‌های انتخابی پوشش مناسبی روی کل آبخوان دارد و هم‌زمان بیش‌ترین اطلاعات مفید در مورد مشخصات کیفی را می‌دهد. همچنین از ایستگاه‌های اضافی به دلیل هزینه‌های موقتی و راه‌اندازی خودداری شده است. روش آنتروپی فرض شده با دو روش Error minimization و K-Means clustering مقایسه شده است و نشان داده شده که روش بهتری است. در مطالعه موردی مربوطه از بین ۷۹ چاه موجود، ۲۰ چاه منتخب با بهترین جواب‌های ممکن برای کیفیت آب بدون نیاز به چاه بیش‌تر بررسی شده‌اند. از روش KNN و ANN برای پرکردن خلأهای اطلاعاتی استفاده شده است. در این مقاله فقط از محتوای اطلاعاتی به‌عنوان یک مقدار در طراحی شبکه پایش استفاده شده است. در صورتی که معیارهای دیگری هم‌چون نیاز انسان، توزیع جمعیت و محدودیت‌های اقتصادی نیز منظور می‌شوند که این مسأله را پیچیده می‌کند. Janatrostami & Salahi (2020) برای

با تحلیل خوشه‌ای سلسله مراتبی براساس فاصله ماهالانویس^۵ پیشنهاد کرده‌اند. این روش برای داده‌های کیفی آب جمع‌آوری شده از آب‌های ساحلی دریاهای بوهای و زرد چین استفاده شده است و برای ارزیابی این روش، داده‌های کیفی با استفاده از فاصله اقلیدسی نیز خوشه‌بندی کرده‌اند و نتایج قابل‌پذیرش بوده است، اما نتایج استفاده از فاصله ماهالانویس مؤثرتر واقع شده است. Alilou *et al.* (2018) یک روش عملی برای اختصاص نقاط نمونه‌گیری در آینده و حال برای منابع آلوده‌کننده غیرنقطه‌ای در یک ناحیه ارائه کرده‌اند. برای دستیابی به این هدف، از روش طول مخلوط رودخانه (RML)^۶ برای پیشنهاد نقاط نمونه‌گیری استفاده شده است. هم‌چنین قبل از انتخاب مناسب‌ترین مکان‌ها برای یک سیستم رودخانه با استفاده از فرایند شبکه تحلیلی (ANP)^۷ طبقه‌بندی اهمیت هر نقطه نمونه‌گیری مشخص شده است. به‌علاوه یک مدل زنجیره‌ای اتوماتا مارکوف سلولی یکپارچه^۸ برای شبیه‌سازی تغییرات آینده منابع غیرنقطه‌ای صورت گرفته است. سرانجام با در نظر گرفتن فعالیت‌های اساسی از طریق نقشه‌برداری از زمین، ارزش سلسله مراتبی (Hierarchy)، ارزش آلودگی احتمالی منابع غیر نقطه‌ای و کمبود بودجه در منطقه مورد مطالعه، هفت نقطه نمونه‌برداری برای حال و آینده مشخص شده است. در کل ۱۵ نقطه نمونه‌گیری شناخته شده که براساس محدودیت‌های مالی، سیستم رودخانه و مدل‌کردن تغییرات ممکن در کاربری اراضی در آینده، هفت نقطه نمونه‌گیری به‌عنوان مناسب‌ترین ایستگاه‌های کنترل پایش کیفی در سیستم رودخانه انتخاب شدند و هم‌چنین دو ایستگاه نیز برای یک شبکه پایش کیفی قوی در صورت توسعه شبکه در اولویت دوم قرار گرفته‌اند. Boroumand *et al.* (2018) یک روش مبتنی بر نیم‌وارپانس-انتقال اطلاعات^۹ برای طراحی یک شبکه بهینه پایش مواد مغذی

برای هر گروه محاسبه کرده‌اند و مکان‌های ایستگاه‌هایی به‌عنوان ایستگاه‌های اضافی مشخص کرده‌اند. نتیجه این رویکرد به تقویت شبکه بارش موجود کمک به‌سزایی کرده است. (Komasi & Goudarzi (2021) دو سناریو برای بهینه‌سازی شبکه پایش آب زیرزمینی ارائه داده‌اند. سناریوی اول شبکه پایش را طراحی می‌کند و سناریوی دوم چاه‌های پایش بهینه را از بین چاه‌های موجود انتخاب می‌کند. در سناریوی اول شبکه پایش توسط الگوریتم ژنتیک غیرغالب و در سناریوی دوم شبکه پایش بهینه توسط تئوری آنتروپی و محاسبه آنتروپی هر یک از ۲۲ چاه مشخص می‌شود. در نتیجه سناریوی اول یک شبکه پایش با ۱۲ چاه و مقدار میانگین مربعات خطای ۰/۶۱ و کاهش ۴۰ درصدی تعداد چاه‌های شبکه پایش موجود را به‌همراه داشته است. سناریوی دوم پس از محاسبه آنتروپی و براساس آن یک شبکه پایش ۱۱ چاهی نتیجه داده است. (Rajaei et al. (2021) توزیع مکانی ایستگاه‌های نمونه‌برداری کیفیت آب سیستم رودخانه‌ای با استفاده از تئوری آنتروپی اطلاعات در حوضه آبریز لواسانات و تهران-کرج با ۱۱ ایستگاه نمونه‌برداری و برای ۱۲ متغیر کیفی و کمی آب در یک دوره آماری ۲۱ ساله، بهینه‌سازی کرده‌اند. نمودار توزیع فاصله بین ایستگاه‌ها و مقدار انتقال اطلاعات را برای هر متغیر مورد مطالعه ترسیم کرده‌اند و با استفاده از روش تصمیم‌گیری چندمعیاره تحلیل سلسله‌مراتبی و با توجه به مصارف آب کشاورزی و شرب، وزن همه متغیرها را تعیین کرده‌اند. بدین ترتیب فاصله بهینه بین ایستگاه‌های پایش به مقدار ۱۴/۱ کیلومتر و هم‌چنین مقدار فاصله بهینه از روش تصمیم‌گیری چندمعیاره تحلیل سلسله‌مراتبی فازی نیز محاسبه کرده‌اند که برابر ۱۳/۸۳ کیلومتر بوده است. سپس با استفاده از این نتایج و استفاده از شاخص‌های دیگری هم‌چون آنتروپی شانون و واریانس،

دستیابی به شبکه پایش کیفی آب زیرزمینی در استان گیلان از الگوریتم بهینه‌سازی ژنتیک استفاده کرده‌اند و هم‌زمان دو تابع هدف متضاد حل کرده‌اند. هدف اول حداکثرکردن تطابق بین توزیع‌های ^{13}EC محاسبه‌شده در شبکه پایش موجود و شبکه جدید با استفاده از کارایی مدل نش-ساتکلیف^{۱۴} بوده و هدف دوم با در نظرگرفتن محدودیت مربوط به هزینه، حداقل کردن تعداد چاه‌های پایش در شبکه جدید بوده است. در شبکه پایش جدید نقاط نمونه‌برداری در مناطق با آلودگی کم‌تر حذف شده و در مناطق با آلودگی بیش‌تر اضافه شده است و در کل تعداد چاه‌های پایش با توجه به توزیع مکانی مقادیر EC در شبکه پایش جدید کاهش یافته است. (Taheri et al. (2020) یک فرایند تحلیل سلسله‌مراتبی^{۱۵} که یک برنامه تصمیم‌گیری چندمعیاره است، برای طراحی یک شبکه پایش کیفی آب زیرزمینی برای یک سفره آب زیرزمینی آبرفتی واقع در محدوده مطالعاتی اسلام‌آباد غرب- کرمانشاه ارائه کرده‌اند. مدل پیشنهادی از هشت معیار اولیه استفاده کرده است و تعداد ۵۹ چاه از ۲۵۴ چاه موجود را به‌عنوان چاه‌های پایش بهینه نتیجه داده است. با در نظرگرفتن پنج زیرمعیار نظیر شرایط فیزیکی چاه‌ها، پمپ‌ها، مجوز استفاده از مالک چاه‌ها و ... که زیر نظر کارشناسان تعیین شده است (در غربالگری دوم چاه‌ها) تعداد چاه‌ها به ۱۳ چاه کاهش یافته است و اهداف سازمان مدیریت آب منطقه‌ای را برآورد کرده است. (Li et al. (2021) یک رویکرد مبتنی بر آنتروپی و کاپولا^{۱۶} برای گسترش شبکه پایش بارش با هدف افزودن ایستگاه‌ها در مناطق غیرقابل سنجش با ارزش پایش بالا و تخمین از طریق محتوای اطلاعات و افزونگی، ارائه کرده‌اند. این رویکرد برای شبکه‌ای در حوضه دریاچه‌ای در چین اعمال شده است. ایستگاه‌های مجاور این دریاچه براساس چندضلعی تیسن^{۱۷} گروه‌بندی شده‌اند، ارزش پایش را

روش‌های احتمالاتی، همبستگی مکانی باقیمانده‌های $\varepsilon(X) = F(X) - f(X)$ به وسیله واریانس فواصل بین نقاط تحلیل می‌شوند. درون‌یابی به وسیله کمینه‌کردن واریانس تخمین انجام می‌گیرد. شکل عمومی رابطه درون‌یابی خطی در یک نقطه فاقد اندازه‌گیری به صورت رابطه (۲) می‌باشد:

$$\hat{Z} = \sum_{i=1}^n W_i \cdot Z_i \quad (\text{رابطه ۲})$$

که \hat{Z} مقدار برآوردی متغیر در نقطه فاقد اندازه‌گیری با مختصات معلوم، W_i وزن نقطه i ام مشاهده‌ای (یا اندازه‌گیری شده)، Z_i مقدار مشاهده شده متغیر در نقطه i (با مختصات معلوم) و n تعداد نقاط مشاهده‌ای می‌باشد به نحوی که $i=1, 2, 3, \dots, n$.

تفاوت روش‌های مختلف درون‌یابی در نحوه محاسبه وزن‌ها (W) می‌باشد. برآورد مقادیر مجهول بر این اصل استوار است که مقادیر معلوم که در اطراف و همسایگی نقطه فاقد اندازه‌گیری (مجهول) قرار دارند، همبستگی بیش‌تر با مقدار متغیر در نقطه مجهول نسبت به نقاط دورتر دارند و لذا انتظار می‌رود وزن نقاط نزدیک‌تر بیش‌تر باشد.

روش زمین‌آماري (کریجینگ)^{۱۸}

زمین‌آمار شاخه‌ای از آمار است که بر مجموعه داده‌های مکانی یا مکانی-زمانی تمرکز دارد و مجموعه‌ای از ابزارهای آماری را با اختلاط ابعاد مکانی و زمانی مشاهدات در پردازش داده‌ها فراهم می‌آورد. این تکنیک نقاط مجهول را براساس خود همبستگی بین نقاط اندازه‌گیری شده و ساختار فضایی آن‌ها تخمین می‌کند. از روش‌های مختلف زمین‌آمار می‌توان کریجینگ و تئوری بی‌زین را نام برد که در این پژوهش از روش کریجینگ استفاده شده است. این روش یک روش تخمین

سناریوهای مختلفی جهت حذف یا افزایش تعداد ایستگاه‌های نمونه‌برداری معرفی کرده‌اند.

با توجه به ارائه مروری بر چندی از پژوهش‌های انجام‌شده، نتیجه گرفته می‌شود که برای طراحی و بازنگری شبکه‌های پایش موجود از روش‌های آماری، زمین‌آماري و شبیه‌سازی-بهینه‌سازی و گاهی ترکیب این روش‌ها استفاده شده است و تا به حال ترکیبی از روش‌های خوشه‌بندی، زمین‌آمار و تئوری آنتروپی برای بازنگری شبکه پایش کیفی آب زیرزمینی استفاده نشده است. از این‌رو، هدف از این پژوهش بازنگری شبکه پایش کیفی آب زیرزمینی با ترکیب این سه روش می‌باشد.

مواد و روش‌ها

در این بخش به روش‌های استفاده شده در این پژوهش پرداخته شده است. در ادامه روش درون‌یابی کریجینگ، روش نزدیک‌ترین همسایگی، تحلیل خوشه‌ای و تئوری آنتروپی ارائه شده‌اند.

روش درون‌یابی

درون‌یاب‌های آماری به دسته‌ای از روش‌ها گفته می‌شود که بر پایه ویژگی‌های آماری سری مکانی متغیر موردنظر، نظیر میانگین و انحراف معیار استوار باشد. یک فرایند آماری می‌تواند به‌عنوان فرایندی که دارای یک بخش دارای ساختار یا قطعی و یک بخش با همبستگی مکانی تصادفی است، تعریف شود. در این حالت بخش دارای ساختار یا دارای روند، با برازش یک تابع خطی یا غیرخطی مناسب تعیین می‌شود؛ بنابراین:

$$F(X) = f(X) + \varepsilon(X) \quad (\text{رابطه ۱})$$

به‌طوری‌که $F(X)$ فرایند آماری، $f(X)$ بخش دارای ساختار و $\varepsilon(X)$ معرف جزء تصادفی می‌باشد. در

چندمتغیره می‌باشد که هدف آن دسته‌بندی داده‌ها براساس تشابه آن‌ها می‌باشد. نقاط معرف داده‌ها در هر دسته (خوشه) باید تا حد امکان مشابه و معرف اعضای هر خوشه و نقاط معرف داده‌ها در خوشه‌های مختلف در حد امکان متفاوت باشند. الگوریتم خوشه‌بندی K-means یک الگوریتم خوشه‌بندی افزایشی است که به دلیل کارایی خود، در خوشه‌بندی مجموعه‌های بزرگ داده‌ها با ویژگی‌های عددی مشهور است. این الگوریتم به دلیل محاسبات ساده و بازدهی مناسب، کاربرد فراوانی در مطالعات مرتبط با منطقه‌بندی دارد (Ahani & Mousavi, 2014). این الگوریتم روشی برای افزایش n داده (x_1, x_2, \dots, x_n) به K خوشه با این مشخصات که داده‌های هر خوشه بیش‌ترین تشابه را با هم و بیش‌ترین تفاوت را با داده‌های خوشه‌های دیگر داشته باشد، به کار می‌رود. این روش یک روند تکراری برای کمینه‌کردن جمع مربعات فاصله هر عضو از هر خوشه تا مرکز هر خوشه $c_j()$ را دنبال می‌کند (Alizadeh et al., 2018). الگوریتم K-means به معیار تشابه (فاصله) و تعداد خوشه‌ها حساس است. در این پژوهش همانند بیش‌تر پژوهش‌های انجام‌شده از فاصله اقلیدسی به‌عنوان معیار تشابه استفاده شده است (Xiong et al., 2006)، که از رابطه (۴) محاسبه می‌شود.

$$d_{euc}(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (\text{رابطه ۴})$$

یک خوشه‌بندی بی‌نقص به مشخص شدن تعداد مناسب خوشه‌ها نیاز دارد که این تعداد با استفاده از معیارهایی سنجیده می‌شود. معیارهایی مانند مجموع مربعات درون خوشه‌ای و معیار سیلوئت (نیمرخ) دو معیار پرکاربرد برای ارزیابی تعداد خوشه‌ها استفاده می‌شوند (Xiong et al., 2006). در این پژوهش از روش سیلوئت، تعداد بهینه خوشه‌ها محاسبه شده است که

زمین‌آمار بر پایه منطق میانگین متحرک وزن‌دار برای درون‌یابی داده‌ها می‌باشد که درون‌یابی داده‌ها را براساس واریانس فضای که آن را تابعی از فاصله می‌شناسد، انجام می‌دهد. تخمین یک مقدار نمونه‌برداری نشده Z در یک مکان مشخص، با استفاده از یک رابطه‌ی خطی به صورت رابطه (۳) محاسبه می‌شود.

$$\hat{Z}(x_0) = \sum_{i=1}^N \omega_i \times Z(x_i) \quad (\text{رابطه ۳})$$

در این معادله، $Z(x_i)$ مقادیر اندازه‌گیری‌شده برای نمونه i ام، ω_i وزن نمونه i ام، x_0 محل تخمین و N تعداد مقادیر اندازه‌گیری‌شده می‌باشد. در روش کریجینگ هر نمونه معلوم در تخمین نقطه مجهول، بستگی کامل به ساختار فضایی محیط مربوط دارد. درحالی‌که در روش‌های دیگر، وزن‌ها فقط به یک مشخصه هندسی مانند فاصله بستگی دارد و با تغییر ساختار فضایی نمونه‌ها، تغییری نمی‌کند و با ضعیف شدن ساختار فضایی، نقش نمونه‌ها کم‌تر می‌شود. تا آنجا که وزن تمام نمونه‌ها برابر خواهد شد. به عبارت دیگر، دامنه تأثیر متغیر معلوم بر مجهول به حداکثر و حداقل فاصله نمونه‌ها از هم بستگی دارد، لذا در استفاده از این روش باید به توزیع فضایی نمونه‌ها و دامنه تأثیر آن‌ها توجه شود. برای برآورد مقادیر براساس کریجینگ روش‌های مختلفی وجود دارد که در این پژوهش از روش کریجینگ عمومی استفاده شده است.

تحلیل خوشه‌ای

در مواردی که تعداد متغیرها، مشاهده‌ها یا داده‌های اندازه‌گیری‌شده هر محدوده مطالعاتی بزرگ باشد، از تحلیل‌های آماری چندمتغیره استفاده می‌شود (Hosseinmardani et al., 2014). تحلیل خوشه‌ای یکی از زیرمجموعه‌های علم داده‌کاوی و از روش‌های آماری

$$H(x) = - \sum_{i=1}^n p(x_i) \ln p(x_i) \quad (\text{رابطه ۵})$$

که در این معادله $p(x_i)$ احتمال وقوع هر رخداد، $\ln p(x_i)$ لگاریتم طبیعی احتمال وقوع هر رخداد و $H(x)$ آنترپی هر رخداد می‌باشد (Masumi & Kerachian, 2008).

منطقه مورد مطالعه

منطقه مطالعاتی، محدوده مطالعاتی تهران-کرج که در شمال شرق حوضه آبریز دریاچه نمک و در جنوب دامنه‌های البرز میانی می‌باشد و به این دلیل که رشته‌کوه البرز این منطقه را از دریای خزر جدا می‌سازد، دریای خزر تأثیر کمی روی این محدوده مطالعاتی می‌گذارد. این محدوده مطالعاتی با مساحت ۲۷۲/۳ کیلومتر مربع بین عرض‌های ۳۵,۱۱ تا ۳۶,۱۸ درجه شمالی و ۵۰,۷۲ تا ۵۱,۷۰ درجه شرقی می‌باشد. کلان‌شهرهای تهران و کرج و نیز شهرهایی مانند پرنده، اسلام‌شهر و ... از مهم‌ترین مراکز مسکونی این محدوده می‌باشند. در شکل (۱) کاربری اراضی محدوده مطالعاتی و همچنین موقعیت شهرهای واقع در این محدوده قابل مشاهده است. طبق این نقشه در قسمت مرکزی مناطق شهری و زمین‌های کشاورزی، در قسمت جنوبی زمین‌های بایر، در شرق و یک نوار باریک در قسمت شمالی شهر تهران زمین‌های با مراتع ضعیف و قسمت شمالی محدوده مطالعاتی نیز مراتع با درجه متوسط را شامل می‌شوند. بلندترین ارتفاعات محدوده مورد مطالعه در شمال حدوداً ۴۳۳۲ متر و کم‌ترین در جنوب به ۸۱۶ متر ارتفاع از سطح دریاها می‌رسد.

داده‌های مربوط به کیفیت از سایت مطالعات منابع آب ایران تهیه شده و به بررسی چاه‌های موجود در محدوده مطالعاتی تهران-کرج پرداخته شده است.

به‌طور کلی کیفیت خوشه‌بندی را ارزیابی می‌کند. هرچه مقدار متوسط سیلوئت بزرگ‌تر باشد خوشه‌بندی بهتر است. روش میانگین سیلوئت، میانگین مقادیر سیلوئت را برای مقادیر مختلف تعداد خوشه محاسبه می‌کند، سپس نمودار میانگین مقادیر سیلوئت-تعداد خوشه رسم می‌شود و تعداد بهینه از روی نمودار، عددی است که مقدار میانگین سیلوئت را در محدوده مشخصی از مقادیر ممکن برای K (تعداد خوشه) به حداکثر برساند.

تئوری آنترپی

این تئوری ملاکی برای سنجش بی‌نظمی در یک سیستم یا میزان اطلاعات موجود در یک فرایند است یعنی هرچه قدر میزان بی‌نظمی یک سیستم بیش‌تر باشد گفته می‌شود آنترپی سیستم بیش‌تر است. وقایع با احتمال وقوع زیاد اطلاعات کم‌تری در اختیار می‌گذارند و برعکس هرچه قدر احتمال وقوع یک رخداد کم‌تر باشد، اطلاعات حاصل از آن بیش‌تر است (Shanon, 1948). در یک تقسیم‌بندی آنترپی به دو دسته پیوسته و گسسته تقسیم می‌شود. در آنترپی پیوسته فرض بر این است که توزیع احتمالاتی متغیرها، از توزیع نرمال یا لوگ‌نرمال پیروی می‌کند (Singh, 1997). در حالت گسسته با توجه به بازه تغییرات مقادیر متغیرها، اطلاعات موجود گسسته‌سازی شده مقادیر احتمال رخداد در هر حالت محاسبه می‌شود. Mogheir & Singh (2006) نشان داد که تابع توزیع احتمال بسیاری از متغیرهای کمی و کیفی در سامانه‌های منابع آب از توزیع‌های نرمال یا لوگ‌نرمال تبعیت نمی‌کنند. آنترپی گسسته راهی برای اصلاح این نقیصه مهم در کاربرد تئوری آنترپی در مسائل مربوط به آب است که در این پژوهش نیز از آنترپی گسسته استفاده شده است. معادله‌ای که شانون به‌عنوان تعریف آنترپی ارائه کرد نیز به‌صورت رابطه (۵) می‌باشد. اگر پایه لگاریتم، ۱۰ باشد واحد آنترپی، هارتلی می‌باشد.

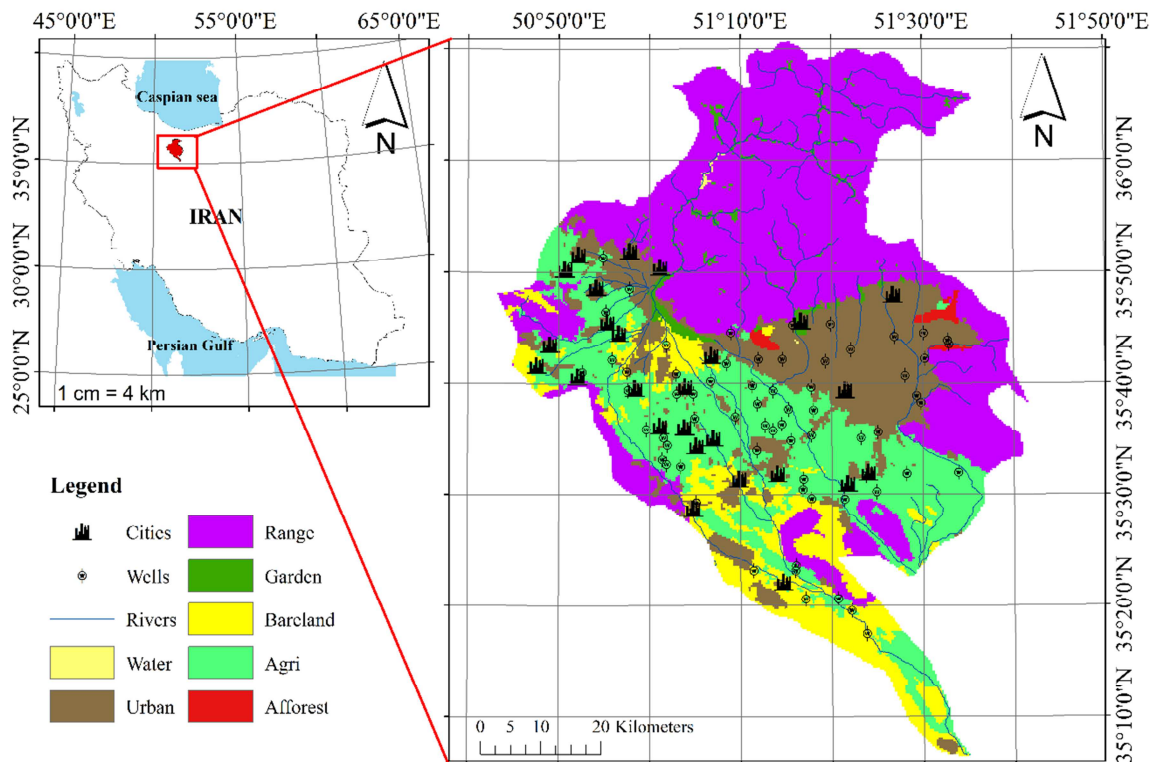


Figure 1. Geographical location of the Tehran-Karaj study area, the geographical location of wells and Landuse

۱۳۹۷ (سالی دو بار و برحسب میلی‌گرم بر لیتر) در ۳۲ دوره، به‌عنوان داده‌های ورودی به مدل پیشنهادی استفاده شده است.

بحث و نتایج

محاسبه مقدار آنتروپی هر چاه

براساس رابطه (۵) برای محاسبه آنتروپی هر چاه به تابع توزیع احتمال ۱۹ مناسب داده‌ها نیاز است. یعنی این‌که داده‌های هر دوره زمانی از چه تابع توزیع احتمالی پیروی می‌کنند. برای یافتن این تابع توزیع احتمال از نرم‌افزار EasyFit استفاده شده است و برای ۳۲ دوره زمانی بهترین توابع توزیع احتمال به داده‌ها برازش داده شده است. به این صورت که معیارهای نکویی برازش را برای هر تابع توزیعی که برازش داده، به‌دست می‌دهد و توابع توزیع احتمال را براساس این معیارها اولویت‌بندی می‌کند.

تعداد چاه‌های محدوده مطالعاتی تهران-کرج ۶۷ چاه و داده‌های کیفی موجود است. با توجه به ضعف فراوان آمار دریافتی، کوتاه‌بودن طول دوره‌های آماری موجود، وجود خلأهای آماری، عدم وجود دقت اندازه‌گیری و هم‌زمان نبودن طول و زمان شروع داده‌های چاه‌های مختلف، لزوم استفاده از روش‌های آماری برای بازسازی و پرکردن خلأهای اطلاعاتی را ایجاب کرده است. از متغیرهای کیفی نمونه‌برداری شده در سطح محدوده مطالعاتی، متغیر TDS (کل مواد جامد محلول) به این دلیل که از لحاظ آماری نسبت به متغیرهای دیگر وضعیت مناسب‌تری داشته است، هم‌چنین به دلیل فعالیت‌های زیاد کشاورزی در مرکز محدوده مطالعاتی و زه‌آب‌های کشاورزی که باعث افزایش TDS در آبخوان شده‌اند، TDS به‌عنوان متغیر مورد مطالعه کیفیت آب زیرزمینی انتخاب شده است. این داده‌ها برای سال‌های ۱۳۸۲ تا

رویکرد اول

در این رویکرد براساس آنتروپی، چاه‌ها به صورت نزولی مرتب می‌شوند. بعد از مرتب‌کردن نزولی آنتروپی چاه‌ها و مشخص کردن چهار چاه که در مرز آنتروپی $0/4$ ، $0/3$ و $0/2$ که به ترتیب در چاه شماره ۳۱، چاه شماره ۴۱ و چاه شماره ۵۱ قرار گرفته‌اند، با استفاده از ۳۱ چاه اول به عنوان شبکه پایش (۴۵ درصد از چاه‌ها)، TDS باقیمانده چاه‌ها از روش کریجینگ تخمین زده خواهد شد و سپس متوسط درصد خطای تخمین محاسبه خواهد شد. در گام بعدی، با استفاده از ۴۱ چاه اول به عنوان شبکه پایش (۶۰ درصد از چاه‌ها)، TDS باقیمانده چاه‌ها از روش کریجینگ تخمین زده خواهد شد و سپس متوسط درصد خطای تخمین با ۴۱ چاه پایش محاسبه خواهد شد و به همین ترتیب برای ۵۱ چاه اول (۷۵ درصد از چاه‌ها) نیز این روند پیش گرفته می‌شود.

رویکرد دوم

در این رویکرد از روش خوشه‌بندی استفاده شده است و هدف تأثیر این روش در روند پایش شبکه موجود می‌باشد. در این بخش به یافتن تعداد خوشه‌ها، خوشه‌بندی و تخمین TDS چاه‌های باقیمانده در هر خوشه با احتساب ۳۰ درصد، ۴۵ درصد و ۶۰ درصد از چاه‌های خوشه به عنوان چاه پایش و در نهایت رسم نمودار خطای تخمین پرداخته شده است. نمودار ارزیابی سیلوئت تعداد خوشه بهینه، در حالت سه‌متغیره در نظر گرفتن خوشه‌بندی، از نرم‌افزار R خروجی گرفته می‌شود و تحت عنوان شکل (۲) ارائه شده است.

طبق شکل (۲) تعداد خوشه بهینه در انتخاب اول دو خوشه و در انتخاب دوم پنج خوشه را نشان داده است که با توجه به این که ممکن است دو خوشه تأثیر چندانی نداشته باشد، تعداد خوشه پنج، به عنوان تعداد

براساس معیار نکویی برازش کای-دو^{۲۰} برای ۳۲ دوره زمانی، تابع توزیع احتمال ویبول^{۲۱} جزو ۱۰ درصد اول این اولویت‌بندی به دست آمده است. بنابراین با این فرض که تابع توزیع احتمال ویبول تابع توزیع مناسبی برای داده‌ها می‌باشد، روند مدل پیشنهادی پیش گرفته شده است. تابع توزیع احتمال در نرم‌افزار MATLAB محاسبه سپس با وجود ۳۲ دوره زمانی، آنتروپی هر چاه محاسبه شده است. مقادیر به دست آمده این آنتروپی‌ها در جدول (۱) ارائه شده است.

Table 1. Entropy values of observation wells (Hartley)

Well	Entropy	Well	Entropy	Well	Entropy	Well	Entropy
1	0.78	18	0.56	35	0.19	52	0.77
2	0.24	19	0.52	36	0.40	53	0.34
3	0.19	20	0.30	37	0.59	54	0.17
4	0.49	21	0.19	38	0.49	55	0.43
5	0.57	22	0.45	39	0.67	56	0.37
6	0.30	23	0.28	40	0.13	57	0.38
7	0.39	24	0.92	41	0.87	58	0.35
8	0.75	25	0.28	42	0.61	59	0.17
9	0.42	26	0.11	43	0.64	60	0.54
10	0.46	27	0.36	44	0.48	61	0.67
11	0.21	28	0.21	45	0.13	62	0.29
12	0.58	29	0.18	46	0.34	63	0.27
13	0.75	30	0.81	47	0.28	64	0.48
14	0.20	31	0.30	48	0.26	65	0.11
15	0.13	32	0.21	49	0.88	66	0.45
16	0.28	33	0.49	50	0.40	67	0.89
17	0.46	34	0.48	51	0.38		

طبق این جدول و طبق مفهوم آنتروپی، هرچه مقدار آنتروپی هر چاه بزرگ‌تر باشد یعنی آن چاه به دلیل آنتروپی بیش‌تر و محتوای اطلاعاتی کم‌تری که دارد، اولویت بیش‌تری برای پایش دارد.

رویکردهای در نظر گرفته شده

برای طراحی شبکه پایش دو رویکرد ارائه شده است. با مقایسه این دو رویکرد نتیجه حاصل، تعداد و مکان چاه‌های پایش محدوده مطالعاتی تهران- کرج می‌باشد. رویکرد اول بدون خوشه‌بندی و رویکرد دوم با اعمال خوشه‌بندی می‌باشد.

بعد از مشخص‌شدن خوشه‌ها چاه‌های هر خوشه براساس آنتروپی هر چاه به‌صورت نزولی مرتب می‌شوند و ۳۰ درصد، ۴۵ درصد و ۶۰ درصد از چاه‌های هر خوشه در نظر گرفته شده است، به این صورت که ابتدا ۳۰ درصد از هر خوشه را به‌عنوان چاه‌های پایش هر خوشه در نظر گرفته و TDS باقیمانده چاه‌های خوشه را با استفاده از روش کریجینگ در نرم‌افزار R تخمین زده و خطای تخمین برای چاه‌های باقیمانده در هر خوشه محاسبه می‌شود. همین روند برای ۴۵ درصد و ۶۰ درصد داده‌های هر خوشه نیز انجام می‌شود.

خوشه بهینه در رویکرد دوم در نظر گرفته می‌شود. براساس تعداد خوشه‌های به‌دست‌آورده‌شده، خوشه‌بندی با روش K-means صورت می‌گیرد و شکل (۳) این خوشه‌بندی را به‌صورت گرافیکی (خروجی نرم‌افزار R) نشان داده است. محور افقی با عنوان Dim1 و محور عمودی با عنوان Dim2 نشان داده شده است. تابع تجزیه و تحلیل داده‌ها در کد روش K-means، مجموعه اولیه متغیرها را به مجموعه جدیدی از متغیرها تبدیل می‌کند تا نتیجه خوشه‌بندی روی دو محور قابل نمایش باشد (Dim1 و Dim2).

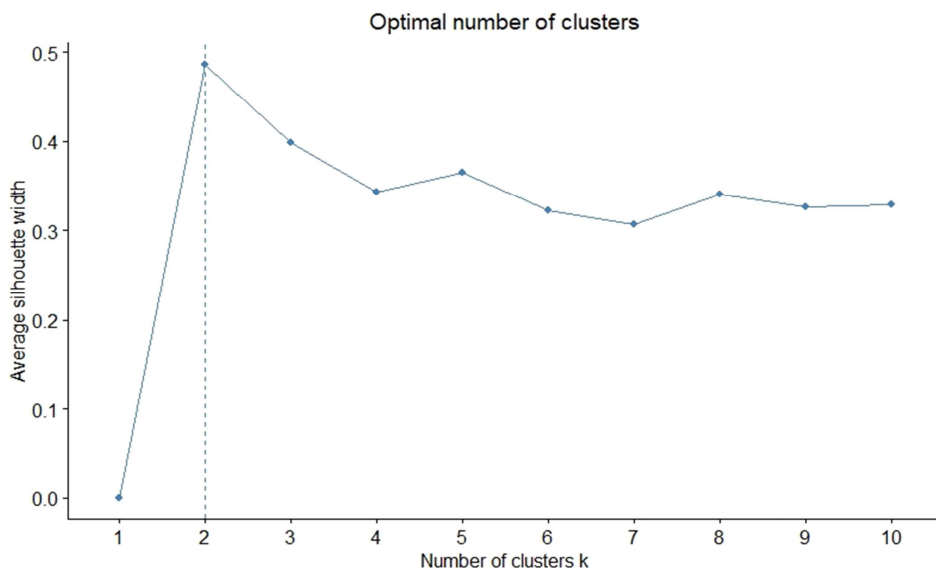


Figure 2. Graph of the optimal number of clusters based on silhouette criteria

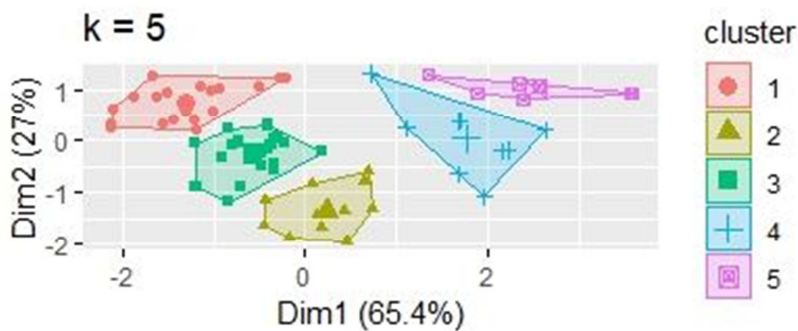


Figure 3. Graph of observation well data clustering

می‌باشد، لذا کم‌ترین متوسط درصد خطای تخمین یعنی ۲۹ درصد مبنای ادامه کار قرار گرفته است. باتوجه به شکل (۴)، حالت با خوشه‌بندی و تعداد خوشه بهینه پنج و تخمین براساس ۳۰ درصد از چاه‌های هر خوشه، مبنای ادامه کار قرار گرفته است. شکل (۵)، نقشه موقعیت جغرافیایی ۳۰ درصد از چاه‌های هر خوشه و چاه‌های دیگر با متوسط درصد خطای تخمین آن‌ها را نمایش داده است.

نمودار جعبه‌ای این حالت در شکل (۶) به تفکیک خوشه قابل مشاهده است. محور افقی این نمودار متوسط درصد خطای تخمین و محور عمودی آن خوشه‌ها را نشان داده است. طبق شکل (۶)، مشاهده می‌شود که خوشه شماره یک بازه متوسط درصد خطای تخمین بیش‌تری نسبت به خوشه‌های دیگر داشته است و در خوشه‌های دو تا چهار نیز میان چاه‌ها خطای در حدود ۳۰ درصد را نشان داده است.

در خوشه دو و چهار ۵۰ درصد از چاه‌ها خطای زیر ۳۰ درصد، در خوشه سه نزدیک به ۷۵ درصد از چاه‌ها خطای زیر ۳۰ درصد و در خوشه پنج نیز بیش‌تر از ۷۵ درصد از چاه‌ها خطای زیر ۳۰ درصد داشته‌اند.

برای مقایسه روش تخمین بدون خوشه‌بندی (رویکرد اول) با روش تخمین با خوشه‌بندی (رویکرد دوم)، نمودار خطای این دو روش قابل مقایسه می‌باشد. شکل (۴) این نمودار را نشان می‌دهد. در این نمودار برای حالت بدون خوشه‌بندی، متوسط درصد خطای تخمین TDS برای چاه‌های باقیمانده (با احتساب ۴۵ درصد، ۶۰ درصد و ۷۵ درصد کل چاه‌های موجود به‌عنوان چاه پایش) نشان داده شده است. هم‌چنین متوسط درصد خطای تخمین TDS برای چاه‌های باقیمانده از هر خوشه (با احتساب ۳۰ درصد، ۴۵ درصد و ۶۰ درصد از چاه‌های هر خوشه به‌عنوان چاه‌های پایش) قابل مشاهده است و از مقایسه این دو حالت بدون خوشه‌بندی و با خوشه‌بندی به‌دست می‌آید که خوشه‌بندی منطقه تأثیر مثبتی بر روی کاهش متوسط درصد خطای تخمین داشته و باعث کاهش ۲۰ درصدی آن شده است. هم‌چنین طبق شکل (۴) قابل مشاهده است که در حالت خوشه‌بندی افزایش تعداد چاه‌ها تأثیر چندانی بر روی کاهش متوسط درصد خطای تخمین نداشته است و تغییرات متوسط درصد خطای تخمین در حد یک درصد

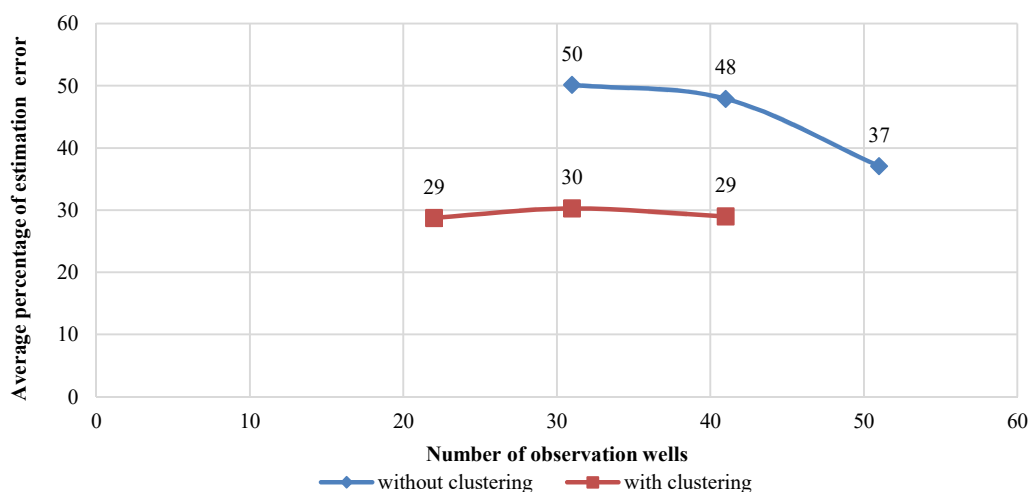


Figure 4. Graph of average percentage of estimation error of existing wells without clustering and with clustering

بازنگری شبکه پایش کیفی آب زیرزمینی با ترکیب روش‌های خوشه‌بندی و زمین‌آمار در محدوده مطالعاتی تهران-کرج

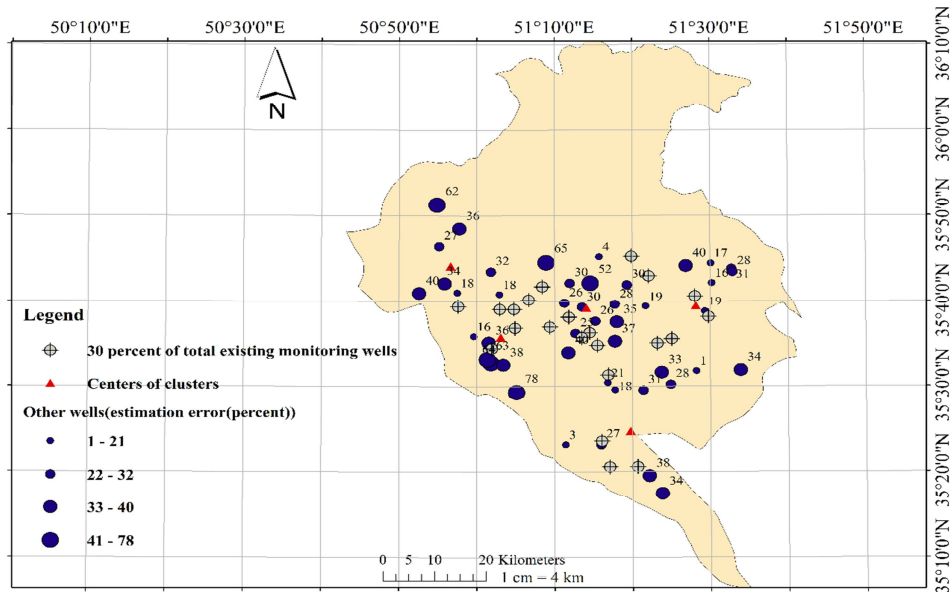


Figure 5. Geographical location of 30 percent of wells and average percentage of estimation error of rest of the wells

خوشه‌ها و محور عمودی نمودار، متوسط درصد خطای تخمین را نشان داده است. طبق این نمودار، نمودارهای میله‌ای آبی‌رنگ متوسط درصد خطای تخمین براساس ۳۰ درصد از چاه‌های هر خوشه، نمودارهای میله‌ای نارنجی‌رنگ متوسط درصد خطای تخمین براساس ۴۵ درصد از چاه‌های هر خوشه و نمودارهای میله‌ای خاکستری‌رنگ متوسط درصد خطای تخمین براساس ۶۰ درصد از چاه‌های هر خوشه را نشان داده است. طبق شکل (۷)، متوسط درصد خطای تخمین براساس ۳۰ درصد از چاه‌های هر خوشه، از دو حالت دیگر (یعنی تخمین براساس ۴۵ درصد و ۶۰ درصد از چاه‌های هر خوشه) کم‌تر بوده است. دلیل این نتیجه می‌تواند این باشد که چاه‌های باقیمانده از هر خوشه با ترکیب ۳۰ درصد از چاه‌های هر خوشه بهتر تخمین زده شده و خطای تخمین کم‌تری ارائه داده است. نمودار میله‌ای زردرنگ نیز متوسط درصد خطای تخمین برای شبکه پایش پیشنهاد شده را نشان داده است. متوسط درصد خطای تخمین شبکه پایش بازنگری شده (متوسط نمودارهای میله‌ای

هم‌چنین مشاهده می‌شود در خوشه‌های یک، دو و پنج، ۵۰ درصد دوم از چاه‌ها (چاه‌های بعد از میانه) تعداد بیش‌تری از چاه‌های هر خوشه را شامل شده‌اند. بنا بر موارد ذکر شده تعدادی از چاه‌ها که خطای بیش‌تر از ۳۰ درصد داشته‌اند به شبکه پایش در نظر گرفته شده در رویکرد دوم (یعنی ۳۰ درصد از چاه‌های هر خوشه به تعداد ۲۲ چاه) اضافه می‌شوند تا همه خوشه‌ها فقط شامل چاه‌های با متوسط درصد خطای تخمین کوچک‌تر از ۳۰ درصد شوند.

طبق شکل (۵) که موقعیت و پراکندگی ۳۰ درصد از چاه‌های موجود و هم‌چنین خطای تخمین باقیمانده چاه‌ها نشان داده شد، از بین چاه‌های باقیمانده، چاه‌های با خطای بالای ۳۰ درصد (۲۲ چاه) به چاه‌های پایش قبلی (تعداد ۲۲ چاه در رویکرد دوم) اضافه می‌شوند. در مجموع ۴۴ چاه به‌عنوان چاه پایش در نظر گرفته می‌شود و خطای تخمین چاه‌های باقی‌مانده محاسبه می‌شود.

شکل (۷)، نمودار این متوسط درصد خطای تخمین را به‌همراه متوسط درصد خطای تخمین رویکرد دوم، به تفکیک خوشه‌ها نشان داده است. محور افقی نمودار،

با توجه به این که بازنگری شبکه پایش با هدف کاهش تعداد چاه‌های پایش موجود انجام گرفته است، لازم است میزان کاهش هزینه نمونه‌برداری نیز که در نتیجه کاهش تعداد چاه‌ها به دست آمده است، ارائه شود. شبکه پایش پیشنهادی در این پژوهش با متوسط درصد خطای تخمین ۱۹ و تعداد ۴۴ حلقه چاه هزینه نمونه‌برداری را نسبت به شبکه پایش فعلی ۳۴ درصد کاهش داده است و هم‌چنین طبق شکل (۹) از پراکندگی مناسبی نیز برخوردار است و از تجمع چاه‌های نزدیک به هم در یک منطقه خاص به دور است.

زردرنگ) ۱۹ درصد می‌باشد. کم‌ترین متوسط درصد خطای رویکرد دوم (متوسط نمودارهای میله‌ای آبی‌رنگ) ۲۹ درصد می‌باشد که نشانگر کاهش ۱۰ درصدی خطای شبکه پایش بازنگری شده می‌باشد.

شکل (۸) موقعیت جغرافیایی و پراکندگی ۴۴ چاه پایش پیشنهاد شده و چاه‌های دیگر با متوسط درصد خطای تخمین آن‌ها را نشان داده است. در این شکل مشاهده می‌شود که چاه‌های باقیمانده خطاهای تخمین کم‌تر از ۳۰ درصد را دارند و چاه‌های با خطای بالاتر از ۳۰ درصد به‌عنوان شبکه پایش پیشنهاد شده‌اند.

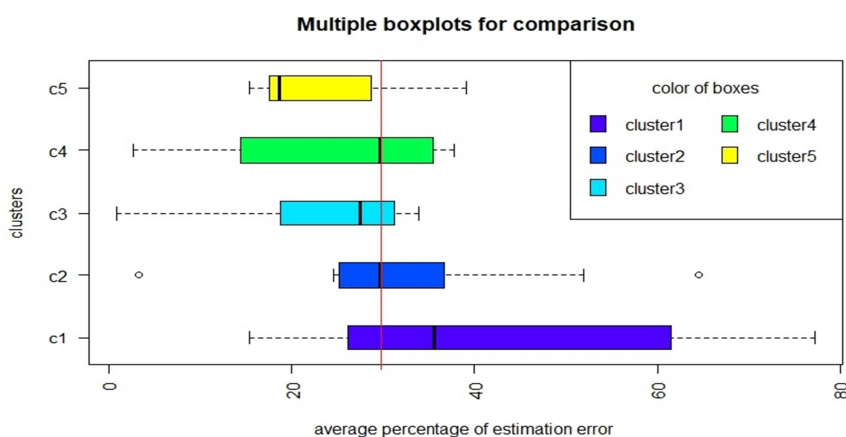


Figure 6. Box Plot of average percentage of estimation error in clusters

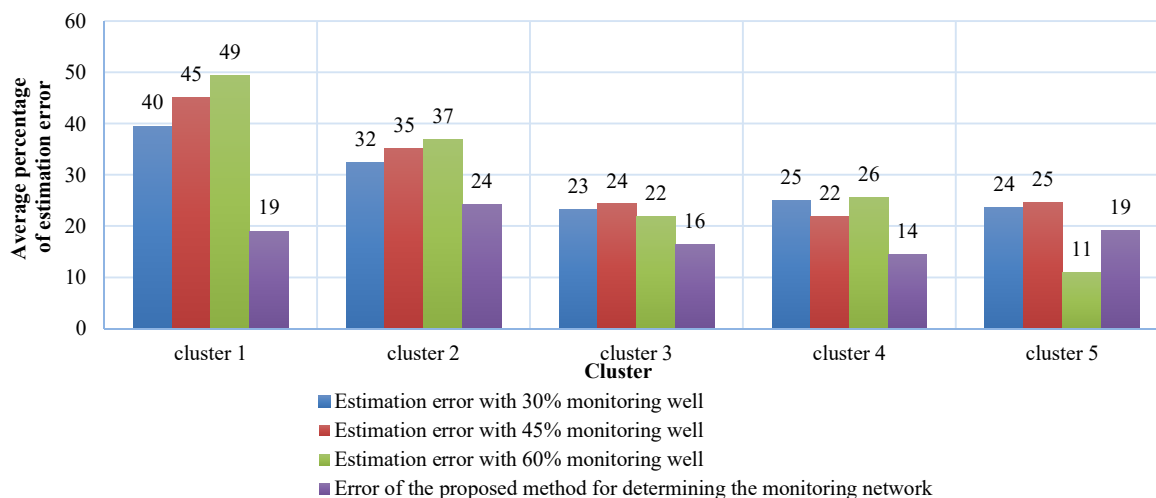


Figure 7. Graph of average percentage of estimation error of wells based on suggested monitoring network

بازنگری شبکه پایش کیفی آب زیرزمینی با ترکیب روش‌های خوشه‌بندی و زمین‌آمار در محدوده مطالعاتی تهران-کرج

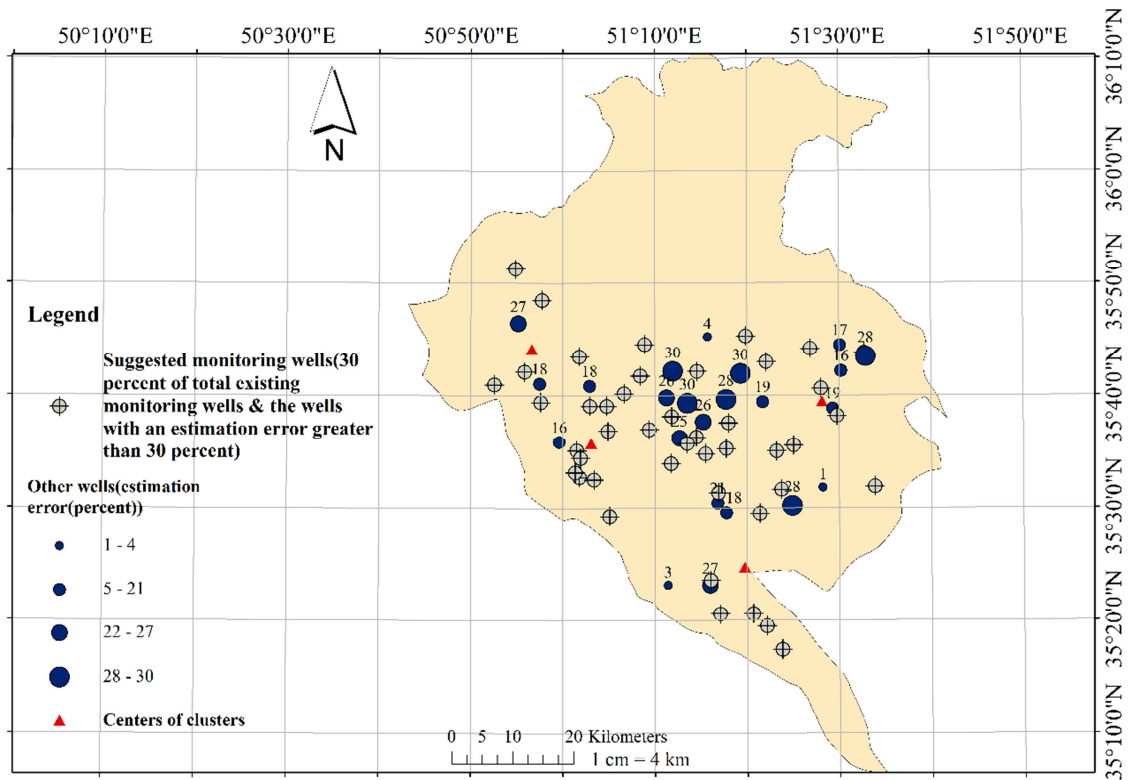


Figure 8. Geographical location of 30 percent of wells, wells with an average percentage of estimation error greater than 30 percent and rest of the wells

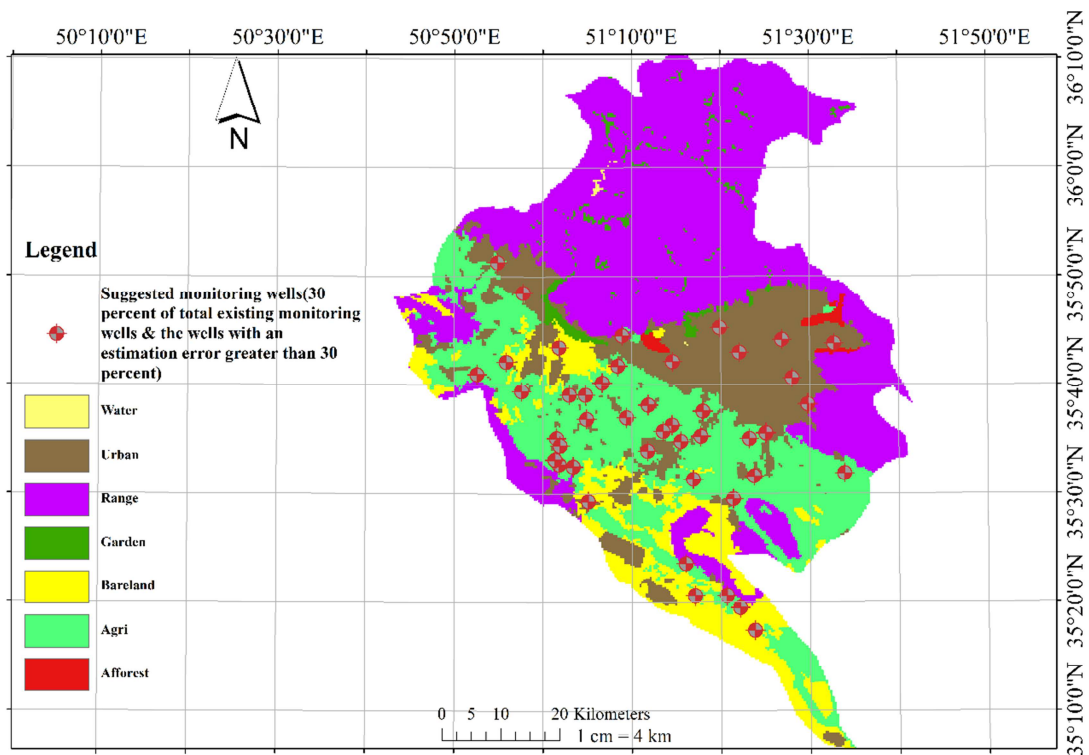


Figure 9. Geographical location and number of suggested monitoring network

مدیریت آب و آبیاری

دوره ۱۲ ■ شماره ۲ ■ تابستان ۱۴۰۱

نتیجه گیری

به‌عنوان چاه پایش رخ داده است مبنای ادامه کار قرار گرفته است. در هر خوشه به‌جز این ۳۰ درصد از چاه‌ها، چاه‌های با خطای تخمین بیش‌تر از ۳۰ درصد به‌عنوان چاه پایش در نظر گرفته شده است و در مجموع ۴۴ چاه به‌عنوان شبکه پایش پیشنهادی ارائه شده است. شبکه پایش نهایی به‌صورت شبکه پایش با متوسط درصد خطای تخمین ۱۹، تعداد ۴۴ حلقه چاه هزینه نمونه‌برداری را نسبت به شبکه پایش فعلی ۳۴ درصد کاهش داده و با پراکندگی مناسب در سطح محدوده مطالعاتی از نظر آنتروپی و TDS متوسط ارائه شده است. به‌صورت خلاصه در این پژوهش تأثیر استفاده از روش‌های تئوری آنتروپی، کریجینگ و خوشه‌بندی بررسی شده است که در محدوده مطالعاتی تهران- کرج باعث کاهش ۳۴ درصدی هزینه پایش با حفظ اطلاعات تولیدی با خطای ۱۹ درصد شده است که می‌توان از این دستاورد برای بازنگری شبکه‌های پایش محدوده‌های مطالعاتی دیگری در کشور استفاده کرد. هم‌چنین ترکیب الگوریتم‌های بهینه‌سازی با روش‌های به‌کاربرده‌شده در این پژوهش، در پژوهش‌های آتی پیشنهاد می‌شود.

پی‌نوشت‌ها

1. Hierarchical cluster analysis
2. Analytical Hierarchy Process
3. Weighted Linear Combination
4. Classification
5. Mahalanobis distance
6. River Mixing Length
7. Analytic Network Process
8. Integrated Cellular Automata-Markov Chain Model
9. Semivariance-Transinformation
10. Total Dissolved Solids
11. Artificial Neural Network
12. K-Nearest Neighbor
13. Electrical conductivity
14. Nash-Sutcliffe
15. Analytical Hierarchical Process
16. Copula
17. Thiessen polygons
18. Krigging
19. Probability Density Function
20. Chi-square
21. Weibull

هدف اصلی این پژوهش، بازنگری شبکه پایش کیفی آب زیرزمینی محدوده مطالعاتی تهران- کرج بوده است که در این راستا پس از مشخص کردن متغیر مورد بررسی (TDS)، با استفاده از تئوری آنتروپی، آنتروپی هر چاه محاسبه شده است. سپس دو رویکرد برای بازنگری شبکه پایش ارائه شده است. در رویکرد اول بدون در نظرگیری خوشه‌بندی چاه‌های موجود، براساس آنتروپی به‌صورت نزولی مرتب می‌شوند و با در نظر گرفتن تعدادی از چاه‌ها طی سه مرحله به‌عنوان چاه‌های پایش، TDS چاه‌های باقیمانده با استفاده از روش کریجینگ به‌دست آمده و خطای تخمین چاه‌های باقیمانده در هر مرحله محاسبه شده است. در رویکرد دوم با استفاده از معیار سیلوئت تعداد خوشه بهینه پنج به‌دست آمده و خوشه‌بندی چاه‌های محدوده مطالعاتی تهران- کرج با روش K-means صورت گرفته است. سپس چاه‌های موجود در هر خوشه براساس آنتروپی به‌صورت نزولی مرتب می‌شوند و با در نظر گرفتن تعدادی از چاه‌های هر خوشه طی سه مرحله به‌عنوان چاه‌های پایش، TDS چاه‌های باقیمانده در هر خوشه با استفاده از روش کریجینگ به‌دست آمده و خطای تخمین چاه‌های باقیمانده در هر خوشه محاسبه شده است. سپس رویکرد اول و دوم مقایسه شده‌اند و کاهش ۱۹ درصدی خطای تخمین در رویکرد دوم (با خوشه‌بندی) نسبت به رویکرد اول (بدون خوشه‌بندی) نشان داده شده است. سپس در رویکرد دوم کم‌ترین خطای تخمین در بین سه مرحله انجام‌شده، انتخاب شده و مبنای روش تعیین تعداد و موقعیت چاه‌های پایش قرار گرفته است. در رویکرد دوم افزایش تعداد چاه‌ها در هر مرحله به‌عنوان چاه‌های پایش تأثیر چندانی بر روی کاهش متوسط درصد خطای تخمین نداشته است و تغییرات متوسط درصد خطای تخمین در حد یک درصد می‌باشد. لذا کم‌ترین متوسط درصد خطای بینی که در تعداد چاه ۳۰ درصد از چاه‌های موجود

- Comparison of New Zealand's national and regional groundwater monitoring programs. *Hydrogeology Journal*, 20(1), 185-200. <https://doi.org/10.1007/s10040-011-0786-2>.
10. Du, X., Shao, F., Wu, S., Zhang, H., & Xu, S. (2017). Water quality assessment with hierarchical cluster analysis based on Mahalanobis distance. *Environmental Monitoring and Assessment*, 189(7). <https://doi.org/10.1007/s10661-017-6035-y>.
 11. Esquivel, J. M., Morales, G. P., & Esteller, M. V. (2015). Groundwater Monitoring Network Design Using GIS and Multicriteria Analysis. *Water Resources Management*, 29(9), 3175-3194. <https://doi.org/10.1007/s11269-015-0989-8>.
 12. Hosseinimrandi, H., Mahdavi, M., Ahmadi, H., Motamedvaziri, B., & Adelpur, A. (2014). Assessment of Groundwater Quality Monitoring Network Using Cluster Analysis, Shib-Kuh Plain, Shur Watershed, Iran. *Journal of Water Resource and Protection*, 06(06), 618-624. <https://doi.org/10.4236/jwarp.2014.66060>.
 13. Janatrostami, S., & Salahi, A. (2020). Design of the optimal groundwater quality monitoring network using a genetic algorithm based optimization approach. *Environmental Sciences*, 18(2), 19-40. (In Persian).
 14. Karamouz, M., Ahmadi, A., & Akhbari, M. (2020). *Groundwater Hydrology: Engineering, Planning, and Management (2nd ed.)*. CRC Press. <https://doi.org/10.1201/9780429265693>.
 15. Komasi, M., & Goudarzi, H. (2021). Multi-objective optimization of groundwater monitoring network using a probability Pareto genetic algorithm and entropy method (case study: Silakhor plain). *Journal of Hydroinformatics*, 23(1), 136-150. <https://doi.org/10.2166/hydro.2020.061>.
 16. Li, H., Wang, D., Singh, V. P., Wang, Y., Wu, J., & Wu, J. (2021). Developing an entropy and copula-based approach for precipitation monitoring network expansion. *Journal of Hydrology*, 598(November 2020), 126366. <https://doi.org/10.1016/j.jhydrol.2021.126366>.
 17. Masoumi, F., & Kerachian, R. (2008). Optimal groundwater monitoring network design using the entropy theory. *J. of Water and Wastewater*, 65, 2-12. (In Persian).
 18. Mogheir, Y., Singh, V. P., & De Lima, J. L. M. P. (2006). Spatial assessment and redesign of a groundwater quality monitoring network using entropy theory, Gaza Strip, Palestine. *Hydrogeology Journal*, 14(5), 700-712.

تعارض منافع

هیچ‌گونه تعارض منافع توسط نویسندگان وجود ندارد.

منابع

1. Ahani, A., & Mousavi Nadoushani, S. S. (2014). Regionalization of Aras Watershed by SOFM. *Iran-Water Resources Research*, 10(3), 88-98. (In Persian).
2. Ahmadi, S.H., & Sedghamiz, A. (2007). Geostatistical Analysis of Spatial and Temporal Variations of Groundwater Level. *Environmental Monitoring and Assessment*, 129, 277-294. <http://dx.doi.org/10.1007/s10661-006-9361-z>
3. Alfonso, L., Lobbrecht, A., & Price, R. (2010). Optimization of water level monitoring network in polder systems using information theory. *Water Resources Research*, 46(1), 1-13. <https://doi.org/10.1029/2009WR008953>.
4. Alfonso, L., He, L., Lobbrecht, A., & Price, R. (2013). Information theory applied to evaluate the discharge monitoring network of the Magdalena River. *Journal of Hydroinformatics*, 15(1), 211-228. <https://doi.org/10.2166/hydro.2012.066>.
5. Alilou, H., Moghaddam Nia, A., Keshtkar, H., Han, D., & Bray, M. (2018). A cost-effective and efficient framework to determine water quality monitoring network locations. *Science of the Total Environment*, 624, 283-293. <https://doi.org/10.1016/j.scitotenv.2017.12.121>.
6. Alizadeh, Z., Yazdi, J., & Moridi, A. (2018). Development of an Entropy Method for Groundwater Quality Monitoring Network Design. *Environmental Processes*, 5(4), 769-788. <https://doi.org/10.1007/s40710-018-0335-2>.
7. Boroumand, A., Rajaei, T., & Masoumi, F. (2018). Semivariance analysis and transinformation entropy for optimal redesigning of nutrients monitoring network in San Francisco bay. *Marine Pollution Bulletin*, 129(2), 689-694. <https://doi.org/10.1016/j.marpolbul.2017.10.057>.
8. Chang, C. L., & Lin, Y. T. (2014). A water quality monitoring network design using fuzzy theory and multiple criteria analysis. *Environmental Monitoring and Assessment*, 186(10), 6459-6469. <https://doi.org/10.1007/s10661-014-3867-6>.
9. Daughney, C. J., Raiber, M., Moreau-Fournier, M., Morgenstern, U., & van der Raaij, R. (2012). Use of hierarchical cluster analysis to assess the representativeness of a baseline groundwater quality monitoring network:

19. Rajaei, T., Masoumi, F., & Ahmadi Siavoshani, F. S. (2021). Optimal location of water quality monitoring stations in river systems by discrete transinformation entropy. *Iranian Journal of Irrigation & Drainage*, 15(2), 295-306. (In Persian).
20. Rezaei, F., Safavi, H. R., & Ahmadi, A. (2013). Groundwater vulnerability assessment using fuzzy logic: a case study in the Zayandehrood aquifers, Iran. *Environmental management*, 51(1), 267-277.
21. Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3), 379-423.
22. Singh, V. P. (1997). The use of entropy in hydrology and water resources. *Hydrological processes*, 11(6), 587-626.
23. Taheri, K., Missimer, T. M., Amini, V., Bahrami, J., & Omidipour, R. (2020). A GIS-expert-based approach for groundwater quality monitoring network design in an alluvial aquifer: a case study and a practical guide. *Environmental Monitoring and Assessment*, 192(11). <https://doi.org/10.1007/s10661-020-08646-y>.
24. Xiong, H., Wu, J., & Chen, J. (2006). *K-means clustering versus validation measures*, 39(2), 779. <https://doi.org/10.1145/1150402.1150503>.