بررسی کیفیت آب رودخانه کرخه با استفاده از آنالیز آماری چند متغیره بر پایه همبستگی و تغییرات داده‌های کیفی

سعید راثی نظامی[۱]، مهرداد نظری‌ها[۲]، اکبر باغوند[۳]، علی مریدی[۴]

مقاله پژوهشی

چکیده

مقدمه: گسترش شهرنشینی و فعالیت‌های صنعتی و کشاورزی در حاشیه رودخانه کرخه و به دنبال آن تخلیه پساب آن‌ها به منابع آب باعث آلودگی آب توسط آلاینده‌های فیزیکی، شیمیایی و بیولوژیکی گردیده است. یکی از مسایل مهم در پایش کیفیت سیستم‌های منابع آب، تشخیص پارامترهای آلاینده اصلی و تفکیک پارامترهای فیزیکی و شیمیایی و رابطه بین آن‌ها می‌باشد. این امر از آن سبب مهم است که پارامترهای آلاینده فیزیکی به طور معمول ناشی از طبیعت و اقلیم حوضه می‌باشد، اما پارامترهای آلاینده شیمیایی از منابع آلاینده انسان ساخت منشأ می‌گیرند. بنابراین اگر همبستگی بین این دو دسته از پارامترهای آلاینده زیاد باشد، می‌توانند به طور طبیعی از منبع یکسانی ناشی شوند. از طرفی یافتن این ارتباطات با روش‌های آزمایشگاهی (مانند روش تفکیک شیمیایی) ممکن است منجر به صرف وقت زیاد و هزینه‌های سنگین شود.

روش‌ها: در تحقیق حاضر، برای یافتن پارامترهای کیفی اصلی مسؤول آلودگی آب رودخانه کرخه از روش آماری آنالیز فاکتور اصلی (Principal component analysis یا PCA) و برای یافتن ارتباط بین پارامترهای فیزیکی و شیمیایی از روش آماری تحلیل همبستگی کانونیک (Canonical correlation analysis یا CCA) استفاده شد. پارامترهای فیزیکی آب رودخانه کرخه شامل: دبی (Q)، کدورت (Turb)، دما (T)، اکسیژن محلول (Dissolved oxygen یا DO)، کل ذرات معلق (Total suspended solids یا TSS) و هدایت الکتریکی (Electerical conductivity یا EC) و پارامترهای شیمیایی آب شامل: pH، اکسیژن‌خواهی بیولوژیکی ۵ روزه (Biological oxygen demand یا BOD_5)، اکسیژن‌خواهی شیمیایی (Chemical oxygen demand یا COD)، فسفر کل (Total phosphor یا TP) و نیتروژن کل (Total nitrogen یا TN) در نمونه‌های جمع‌آوری شده از ۱۸ ایستگاه در طول رودخانه کرخه بود که در چهار نوبت (در سال ۱۳۸۵) اندازه‌گیری شدند.

یافته‌ها: آنالیز PCA نشان داد که تمامی پارامترهای کیفی انتخاب شده برای بررسی کیفیت آب رودخانه کرخه مهم بودند. در آنالیز CCA، ضریب همبستگی برای سه متغیر استاندارد اول، به ترتیب ۰/۸۹۶، ۰/۸۴۸ و ۰/۵۹۳ بودند که نشانگر این بود که Turb، DO و TSS از بین پارامترهای فیزیکی و BOD_5، COD، TP و TN از بین پارامترهای شیمیایی، دارای بیشترین امتیاز در کل متغیرهای کانونیک بودند.

نتیجه‌گیری: بر اساس نتایج به دست آمده از تحلیل همبستگی کانونیک برای داده‌های کیفی مورد بررسی در این تحقیق، مشخص می‌شود که منشأ هر دو دسته پارامترهای آلاینده شیمیایی و فیزیکی آب رودخانه، انسان ساخت می‌باشد. روش‌های آماری معرفی شده می‌توانند به عنوان ابزار پشتیبان، تصمیم‌گیری مناسبی جهت بررسی تغییرات کیفی منابع آب در دسترس در مواردی که حجم زیادی از داده‌های کیفی وجود داشته باشد، توسط تصمیم گیرندگان و مدیران آب کشور مورد استفاده قرار گیرد.

واژه‌های کلیدی: رودخانه کرخه، آنالیز فاکتور اصلی، تحلیل همبستگی کانونیک، متغیرهای کانونیک

ارجاع: راثی نظامی سعید، نظری‌ها مهرداد، باغوند اکبر، مریدی علی. **بررسی کیفیت آب رودخانه کرخه با استفاده از آنالیز آماری چند متغیره بر پایه همبستگی و تغییرات داده‌های کیفی.** مجله تحقیقات نظام سلامت ۱۳۹۱؛ ۸ (۷): ۱۲۹۲-۱۲۸۰.

دریافت مقاله: ۱۳۹۱/۲/۲۳ پذیرش مقاله: ۱۳۹۱/۵/۲۶

۱- دانشجوی دکتری، گروه مهندسی عمران- محیط زیست، دانشکده محیط زیست، دانشگاه تهران، تهران، ایران (نویسنده مسؤول)
Email: saeedrasinezami@yahoo.com

۲- استادیار، گروه مهندسی محیط زیست، دانشکده محیط زیست، دانشگاه تهران، تهران، ایران

۳- دانشیار، گروه مهندسی محیط زیست، دانشکده محیط زیست، دانشگاه تهران، تهران، ایران

۴- استادیار، گروه مهندسی منابع آب، دانشکده مهندسی آب و محیط زیست، دانشگاه صنعت آب و برق شهید عباسپور، تهران، ایران

مقدمه

آلودگی آب توسط آلاینده‌های شیمیایی، فیزیکی و بیولوژیکی در تمام دنیا به عنوان یکی از مشکلات مدیریت منابع آب در نظر گرفته می‌شود (۲، ۱). گذشته از آن، رودخانه‌ها نقش مهمی را در جذب یا انتقال فاضلاب‌های شهری و صنعتی و رواناب‌های سطحی از زمین‌های کشاورزی دارند (۳). بنابراین، ارزیابی کیفی آب رودخانه دارای اهمیت فراوانی است؛ چرا که به طور مستقیم بر سلامت عمومی (از طریق آب آشامیدنی) و زندگی آبی (از طریق آب خام) تأثیر می‌گذارد. در طول دهه‌های گذشته، پایش کیفی آب رودخانه با اندازه‌گیری پارامترهای کیفی متعدد آب بیشتر شده است. با این حال، به دلیل تغییرات مکانی و زمانی در کیفیت آب که اغلب به راحتی قابل تفسیر نیستند، یک برنامه پایش شامل یک تخمین قابل اعتماد از کیفیت آب‌های سطحی ضروری است (۴). عمده منابع آلاینده آب‌های سطحی شامل منابع نقطه‌ای مانند زهکش‌های کشاورزی، پساب‌های شهری و صنعتی و منابع گسترده مانند زه‌آب‌های کشاورزی می‌باشند. با توجه به اهمیت آب و اثراتی که آلودگی آن می‌تواند بر سلامت انسان و اکوسیستم طبیعی داشته باشد، پایش کیفیت آب از اهمیت بسزایی برخوردار است (۵). در بین منابع آب‌های سطحی، رودخانه‌ها به دلیل طول زیاد و عبور از مناطق مختلف شهری و صنعتی از پتانسیل بیشتری برای آلودگی برخوردار هستند. این موضوع در مناطقی که شهرها و صنایع بزرگ در نزدیکی رودخانه‌ها قرار داشته و رودخانه‌ها تأمین کننده اصلی نیازهای آبی می‌باشند، به منظور حفظ سلامت عمومی، بیشتر مورد توجه قرار گرفته است. یکی از مهم‌ترین راه‌حل‌ها جهت کاهش اثرات منفی آلودگی و اتخاذ تدابیر مدیریتی مناسب جهت کنترل کیفی آب، ایجاد شبکه کنترل کیفی با کارایی مناسب بر روی چنین سیستم‌های آبی است (۳).

یکی از مسایل مهم در پایش کیفیت سیستم‌های منابع آب، تشخیص و تفکیک پارامترهای آلاینده فیزیکی و شیمیایی و رابطه بین آن‌ها می‌باشد. این امر از آن سبب مهم می‌باشد که پارامترهای آلاینده فیزیکی به طور معمول ناشی از طبیعت و اقلیم حوضه می‌باشد، اما پارامترهای آلاینده شیمیایی از منابع آلاینده انسان ساخت منشأ می‌گیرند. بنابراین اگر همبستگی بین این دو دسته از پارامترهای آلاینده زیاد باشد، می‌توانند به طور طبیعی مانند یکدیگر و از منبع یکسانی ناشی شوند. از طرف دیگر، یافتن این ارتباطات با روش‌های آزمایشگاهی مانند روش تفکیک شیمیایی هزینه‌های سنگینی در بر دارد. یکی از گزینه‌ها برای کاهش هزینه‌ها، استفاده از روش آماری چند متغیره تحلیل مؤلفه اصلی (Principal Component Analysis یا PCA) برای شناسایی پارامترهای کیفی اصلی و تحلیل همبستگی کانونیک (Canonical correlation analysis یا CCA) برای یافتن ارتباط بین پارامترهای آلاینده فیزیکی و شیمیایی است.

Barnett و Preisendorfer با روش CCA دمای هوا در آمریکا را از دمای سطح اقیانوس و سطح فشار دریا پیش‌بینی نمودند (۶). Statheropoulos و همکاران از این روش برای تخمین رابطه بین دو دسته از اطلاعات مربوط به آلودگی هوا و اطلاعات هواشناسی در یک ایستگاه کنترل آلودگی هوا واقع در شهر آتن استفاده کردند. نتایج تحقیق مذکور ارتباط اصلی بین آلودگی کل و رطوبت هوا در ترکیب با سرعت پایین باد را نتیجه داد (۷). Larson و همکاران از CCA برای آنالیز رابطه بین امواج و پروفیل‌های برداشت شده از منطقه داک در کارولینای شمالی بهره بردند (۸). Ouarda و همکاران از تحلیل همبستگی کانونیک جهت تخمین نوسانات سیل در حوضه‌ای واقع در کانادا استفاده نمودند (۹). Gramotnev و Gramotnev تحلیل همبستگی کانونیک را جهت تعیین ارتباط بین ذرات معلق ناشی از سوخت خودروها و پارامترهای محیطی مانند ترافیک به کار گرفتند (۱۰). رسولی و همکاران با استفاده از تحلیل همبستگی کانونیک الگوی مناسبی برای مدیریت کیفی سیستم‌های منابع آب ارایه نمودند. برای این منظور رودخانه سفیدرود به عنوان محدوده مطالعاتی انتخاب گردید و نتایج بیانگر منشأ انسان‌ساخت پارامترهای آلاینده فیزیکی و شیمیایی بودند (۱۱).

در تحقیقی برای تفسیر بهتر و مشخص‌سازی منابع آلاینده فلزات سنگین، ترکیبات آلی و دیگر پارامترهای فیزیکی رودخانه‌ای در شمال شرقی اسپانیا از PCA در تلفیق

www.mui.ac.ir

با سیستم اطلاعات جغرافیایی استفاده شد (۱۲). Simeonov و همکاران با استفاده از روش‌های آنالیز کلاستر (Cluster analysis)، PCA و رگرسیون خطی چند متغیره به تجزیه و تحلیل و تفسیر ماتریس اطلاعات حجیم و پیچیده داده‌های کیفی آب که در طول ۳ سال اندازه‌گیری از ۲۵ ایستگاه نمونه‌برداری (ماهانه) آب‌های سطحی در شمال یونان به دست آمده بود، پرداختند. نتایج تحقیق آن‌ها حاکی از اجرای موفقیت‌آمیز روش‌های مذکور به منظور ارزیابی کیفی آب‌های سطحی و طراحی ایستگاه‌های نمونه‌برداری بود (۱۳). از دیگر کاربردهای PCA می‌توان به ارزیابی نوسانات زمانی و مکانی کیفیت آب رودخانه گمتی در هند (۱۴) و ارزیابی و مدیریت کیفی آب‌های سطحی (۱۵) اشاره نمود. مطالعات ذکر شده باعث به وجود آمدن دیدگاهی ارزشمند در استفاده از روش PCA در مدیریت محیط زیست شده است.

Gangopadhyay و همکاران تکنیک PCA را برای شناسایی اهمیت چاه‌های پایش پیش‌بینی کننده متغیرهای پویای مرتبط با تراز پنوماتیک در منطقه‌ای در بانکوک تایلند به کار بردند (۱۶). با استفاده از رهیافت‌های PCA و GIS، Terrado و همکاران منابع آلاینده اصلی فلزات سنگین، ترکیبات آلی و پارامترهای فیزیکی- شیمیایی دیگر را در آب‌های سطحی رودخانه Elbo مورد تجزیه و تحلیل قرار دادند (۱۲). Ouyang، PCA را برای شناسایی پارامترهای کیفی مهم آب در ۲۲ ایستگاه واقع در بخش اصلی رودخانه St.Johns سفلی در فلوریدای آمریکا به کار برد (۱۷). نتایج نشانگر این بودند که کل کربن آلی (TOC **یا** Total organic carbon)، کربن آلی محلول (DOC **یا** Dissolved organic carbon)، نیتروژن کل (TN **یا** Total Nitrogen)، نیترات و نیتریت محلول، اورتوفسفات، قلیائیت، شوری، منیزیم و کلسیم، مهم‌ترین پارامترها در ارزیابی متغیرهای کیفی آب رودخانه بودند. Shrestha و Kazama روش PCA را برای ارزیابی تغییرات زمانی و مکانی دسته اطلاعات بزرگ و پیچیده کیفی آب در حوضه آبریز رودخانه Fuji به کار بردند (۳).

چنان که در تحقیقات اشاره شده در فوق ملاحظه می‌گردد، از تکنیک‌های آماری CCA و PCA به طور جداگانه در جهت نیل به اهداف مختلف استفاده گردیده است، ولی در هیچ کدام از این تحقیقات از تلفیق این دو روش آماری برای مدیریت کیفی حوضه رودخانه استفاده نشده است. با در نظر گرفتن تحقیقات فوق، هدف از تحقیق حاضر، تعیین پارامترهای اصلی و بررسی ارتباط بین پارامترهای فیزیکی و شیمیایی رودخانه کرخه با استفاده از آنالیزهای آماری چند متغیره PCA و CCA بر پایه همبستگی و تغییرات پارامترها جهت تدوین یک سیستم پشتیبان تصمیم‌گیری در مدیریت کیفی رودخانه در مواقعی که تصمیم گیرندگان با دسته اطلاعات بزرگ کیفی مواجه هستند، می‌باشد.

واضح است که بررسی و تعیین روابط بین پارامترهای فیزیکی و شیمیایی آب می‌تواند تصمیم گیرندگان را در مدیریت کیفی آب رودخانه کمک کند. پارامترهایی نظیر کدورت (Turb)، کل ذرات معلق (Total suspended solids یا TSS)، کل ذرات محلول (Total dissolved solids یا TDS) و ... از ویژگی‌های فیزیکی رودخانه نشأت می‌گیرند؛ در حالی که اکسیژن‌خواهی شیمیایی (COD یا Chemical oxygen demand) و یون‌های سولفات (SO_4^{-2}) و نیترات (NO_3^-) به طور معمول نشانگر آلودگی آب توسط فعالیت‌های انسانی هستند.

روش‌ها

مطالعه موردی و داده‌ها

حوضه رودخانه کرخه که توسط رودخانه کرخه زهکشی می‌شود، در غرب ایران قرار دارد (شکل ۱). مساحت حوضه ۵۰۰۰۰ کیلومتر مربع بوده و در محدوده مختصات ۲۳′ ۴۶° تا ۱۲′ ۴۹° طول جغرافیایی و ۴۰′ ۳۳° تا ۰۰′ ۳۵° عرض جغرافیایی قرار گرفته است و از لحاظ تقسیمات کشوری در استان‌های همدان، کرمانشاه، ایلام، لرستان و خوزستان واقع شده است. شاخه اصلی رودخانه کارون از رشته کوه‌های زاگرس سرچشمه گرفته و به سمت شمال غربی خلیج فارس جریان می‌یابد. رودخانه کرخه با متوسط آب‌دهی ۱۷۷ متر مکعب بر ثانیه پس از رودخانه‌های دز و کارون، بزرگ‌ترین

www.mui.ac.ir

رودخانه کشور محسوب می‌شود. این رودخانه از مناطق میانی و جنوب غربی رشته کوه‌های زاگرس سرچشمه گرفته و پس از طی مسافتی حدود ۹۰۰ کیلومتر در امتداد شمال به جنوب، سرانجام در مرز مشترک ایران و عراق به مرداب هورالعظیم و هورالهویزه در جنوب غربی استان خوزستان می‌ریزد.

رودخانه کرخه منبع تأمین آب شرب برای نقاط شهری و همچنین آب مورد نیاز فعالیت‌های کشاورزی و صنعتی است. شهرها و صنایع متعدد شامل فولاد، نفت، پتروشیمی، نیشکر، کاغذ و سیمان و همچنین اراضی کشاورزی تحت آبیاری زیاد از آب رودخانه کرخه و سرشاخه‌های آن استفاده نموده و فاضلاب خود را به آن تخلیه می‌کنند. ساخت سدهای متعدد در طول رودخانه کرخه برداشت آب از رودخانه را افزایش داده است. به علاوه، تخلیه فاضلاب‌های شهری، صنعتی و کشاورزی منجر به تخریب جدی کیفیت آب به خصوص در پایین دست صنایع در رودخانه شده است. کمیت و کیفیت رودخانه کرخه نقش مهمی را در توسعه پایدار صنعتی و کشاورزی استان‌های واقع در حوضه آبریز آن ایفا می‌کند. بنابراین پایش کیفی آب و یافتن روش‌هایی برای محدودیت آلودگی آب باید مورد توجه و تأکید قرار گیرد.

در طول یک سال (از سال ۱۳۸۵ تا ۱۳۸۶) داده‌ها (۲۹ پارامتر برای ارزیابی کیفی آب) از ۴۰ ایستگاه پایش کیفی آب در طول رودخانه کرخه و سرشاخه‌های آن جمع‌آوری شده است. به دلیل پیوستگی داده‌ها در اندازه‌گیری، از بین آن‌ها، ۱۱ پارامتر در ۱۸ ایستگاه در تحقیق ما مورد استفاده قرار گرفت (شکل ۱). پارامترهای فیزیکی و شیمیایی انتخاب شده برای آب رودخانه شامل دبی (Q)، کدورت (Turb)، دما (T)، اکسیژن محلول (Dissolved oxygen یا DO)، کل ذرات معلق (TSS)، هدایت الکتریکی (Electerical conductivity یا EC)، pH، اکسیژن‌خواهی بیولوژیکی ۵ روزه (BOD_5 یا Biological oxygen demand)، اکسیژن‌خواهی شیمیایی (COD)، فسفر کل (Total phosphor) (TP) و نیتروژن کل (Total nitrogen یا TN) بودند. داده‌ها دارای توزیع نرمال نبوده و دارای اریب مثبت در توزیع بودند. بنابراین، مقادیر میانگین سالانه برای هر پارامتر مورد استفاده قرار گرفت (۱۷).

در تحقیق حاضر تعیین پارامترهای اصلی کیفی با استفاده از روش آماری چند متغیره تحلیل مؤلفه اصلی (PCA) و یافتن ارتباط بین پارامترهای آلاینده فیزیکی و شیمیایی با استفاده از تحلیل همبستگی کانونیک (CCA) انجام می‌پذیرد.

تحلیل مؤلفه اصلی (PCA)

PCA از روش‌های آماری چند متغیره می‌باشد که می‌توان از آن برای کاهش پیچیدگی تحلیل متغیرهای اولیه مسأله در مواردی که با حجم زیادی از اطلاعات روبرو هستیم و تفسیر بهتر آن‌ها استفاده نمود (۱). با اعمال این روش، متغیرهای اولیه به مؤلفه‌های جدید و مستقل از یکدیگر (با ضریب همبستگی صفر برای هر دو مؤلفه) تبدیل می‌شوند. مؤلفه‌های جدید ایجاد شده ترکیبی خطی از متغیرهای اولیه هستند (۱۸). با استفاده از این تکنیک‌ها، ترکیباتی از P متغیر اولیه X_1، X_2، ...، X_p برای ایجاد P مؤلفه مستقل (معادل با تعداد متغیرهای اولیه مورد استفاده)، Z_1، Z_2، ...، Z_p ایجاد می‌شود. عدم همبستگی بین این مؤلفه‌ها ویژگی مفیدی است؛ چرا که عدم همبستگی به این معنی است که مؤلفه‌ها، جنبه‌های متفاوتی از متغیرهای اولیه را نمایان می‌سازند (۱۹). در این روش به جای استفاده مستقیم از متغیرهای اولیه، ابتدا آن‌ها را به مؤلفه‌های جدیدی تبدیل کرده، سپس از این مؤلفه‌ها به جای متغیرهای اولیه استفاده می‌گردد. همچنین به دلیل آن‌که در تشکیل مؤلفه‌ها از تمام متغیرها استفاده می‌گردد، اطلاعات متغیرهای اولیه با کم‌ترین تلفات توسط مؤلفه‌های حاصل ارایه می‌شود. در این روش‌ها هر مؤلفه اصلی می‌تواند با دنباله زیر مشخص شود:

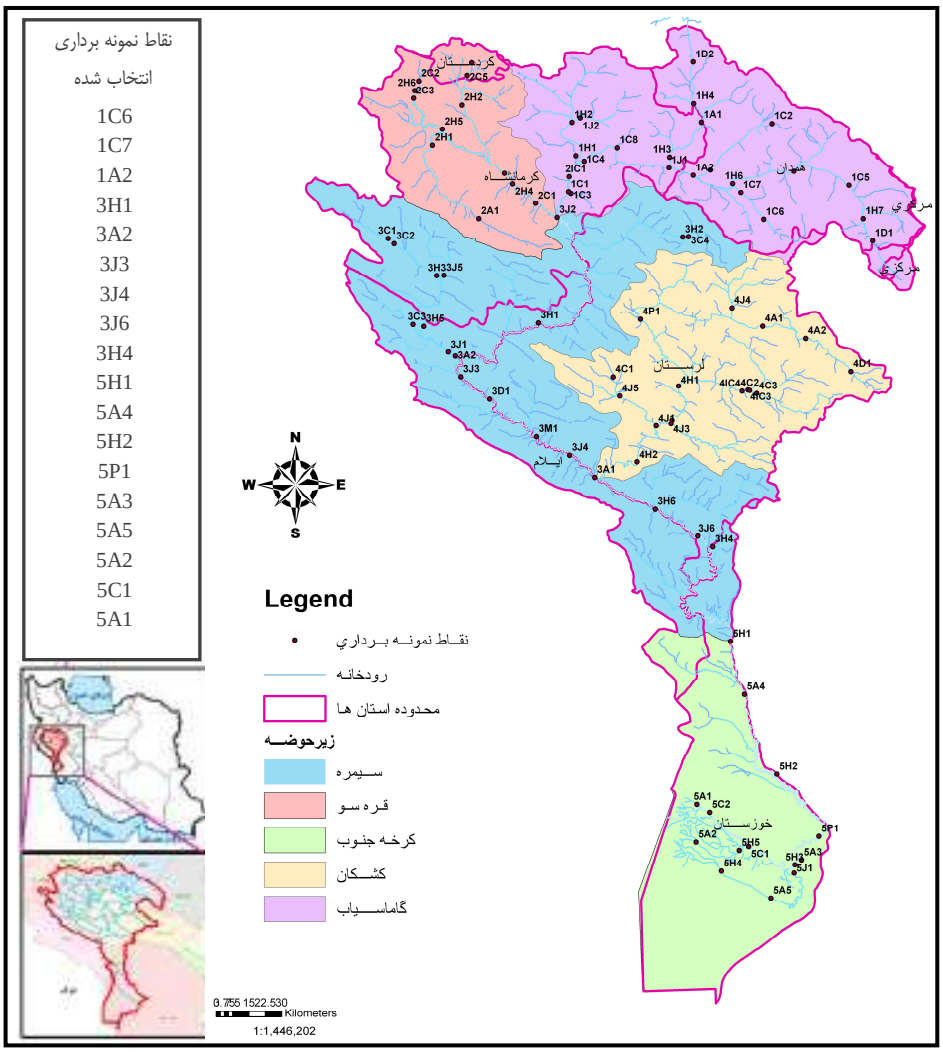
فرمول ۱: $$Z_i = a_{i1}X_1 + a_{i2}X_2 + \cdots + a_{ip}X_p$$

در فرمول ۱، Z_i معرف مؤلفه مورد نظر، a_{ij} ضرایب مربوط به متغیرهای اولیه و X_i نیز متغیرهای اولیه می‌باشند. ضرایب مربوط به متغیرهای اولیه از حل معادله زیر به دست می‌آید.

فرمول ۲: $$\left| R - \lambda I \right| = 0$$

در فرمول ۲، I ماتریس واحد، R ماتریس همبستگی بین متغیرهای اولیه و λ نیز مقادیر ویژه می‌باشند. از مقادیر ویژه، بردارهای ویژه به دست می‌آیند. برای انجام این دو تکنیک مراحل زیر باید انجام شود:

www.mui.ac.ir



شکل ۱: جانمایی ایستگاه‌های نمونه‌برداری در طول رودخانه کرخه (۱۹)

۱- استاندارد نمودن متغیرهای ورودی: در این مرحله داده‌های ورودی به نحوی استاندارد می‌شوند که دارای میانگین صفر و انحراف معیار یک باشند.

۲- محاسبه ماتریس همبستگی برای متغیرهای اولیه: این ماتریس که ماتریسی متقارن می‌باشد، میزان همبستگی بین هر کدام از متغیرهای اولیه مورد استفاده را نشان می‌دهد. در حقیقت مقدار هر کدام از درایه‌های این ماتریس، a_{ij} بیانگر همبستگی بین متغیرهای i و j می‌باشد.

۳- محاسبه مقادیر ویژه λ و بردارهای ویژه مربوط از ماتریس همبستگی: با حل معادله ۲ مقادیر ویژه و بردارهای ویژه معادل هر مقدار ویژه به دست می‌آیند. هر مقدار ویژه با اطلاعات مربوط به آن (بردارهای ویژه) ویژگی‌های یک مؤلفه را نشان می‌دهد. هر مؤلفه نیز درصدی از اطلاعاتی که توسط متغیرهای اولیه بیان می‌شود را در بر می‌گیرد و معادل با بخشی از اطلاعات مسأله است که به صورت عدد و رقم در متغیرهای اولیه نهفته است و چون ماتریس همبستگی ماتریسی متقارن از مرتبه n می‌باشد، مجموع مقادیر ویژه برای این ماتریس نیز برابر با n است. بردارهای ویژه به دست آمده به ازای هر مقدار ویژه نیز، به عنوان ضرایب متغیرهای اولیه در تشکیل مؤلفه مربوط هستند.

اجرای چرخش مناسب بر روی ماتریس ضرایب مؤلفه‌ها: چون در تشکیل هر مؤلفه از تمام متغیرهای اولیه استفاده

www.mui.ac.ir

جدول ۱: پارامترهای کیفی آب سطحی برای رودخانه کرخه

ایستگاه‌ها	پارامترهای فیزیکی و شیمیایی رودخانه										
	Q (مترمکعب بر ثانیه)	Turb. (میلی‌گرم بر لیتر)	T (میلی‌گرم بر لیتر)	DO (میلی‌گرم بر لیتر)	TSS (میلی‌گرم بر لیتر)	EC (میکروموس بر سانتیمتر)	pH -	BOD_5 (میلی‌گرم بر لیتر)	COD (میلی‌گرم بر لیتر)	TP (میلی‌گرم بر لیتر)	TN (میلی‌گرم بر لیتر)
1C6	۱/۷۶	۸۲/۵۰	۱۲/۶۰	۱۰/۷۹	۱۳/۰۰	۴۷۶/۰۰	۷/۷۴	۸/۶۲	۱۶/۳۳	۲/۴۲	۱۲/۹۳
1C7	۲/۱۰	۱۲۵/۷۵	۱۲/۱۸	۸/۹۷	۱۹/۵۲	۵۹۶/۲۵	۷/۷۳	۴/۴۷	۱۴/۰۰	۱/۲۴	۹/۰۶
1A2	۴/۷۰	۵۴/۶۷	۱۱/۸۰	۱۱/۰۹	۳۴/۰۰	۵۵۵/۳۳	۷/۸۷	۷/۴۷	۳۳/۶۷	۲/۶۹	۹/۲۱
3H1	۲۲/۲۲	۴۵۶/۲۵	۱۷/۷۵	۱۳/۶۹	۶۸۰/۷۵	۵۵۲/۵۰	۸/۶۵	۷/۲۰	۲۵/۳۵	۱/۱۵	۱۳/۶۰
3A2	۱۰/۴۹	۱۹۳/۳۳	۲۰/۸۷	۱۰/۵۱	۱۷/۳۳	۹۵۲/۳۳	۷/۸۰	۶/۱۷	۱۹/۲۷	۰/۳۳	۱/۸۱
3J3	۳۷/۲۰	۲۶۹/۲۵	۱۷/۸۸	۱۰/۹۰	۳۰/۷۶	۶۹۵/۰۰	۸/۵۸	۵/۵۰	۱۲/۸۷	۰/۸۸	۱۱/۳۹
3J4	۳۵/۱۹	۳۰۳/۵۰	۱۹/۱۵	۱۰/۸۸	۱۴۳/۰۱	۹۶۷/۵۰	۸/۴۵	۱۴/۸۵	۲۷/۱۵	۰/۳۶	۱۳/۳۱
3J6	۲۴/۳۳	۴۲۲/۲۵	۱۹/۰۳	۱۰/۹۰	۳۱۱/۷۶	۱۵۳۸/۲۵	۸/۲۸	۹/۲۱	۲۹/۳۸	۰/۶۲	۳/۷۶
3H4	۲/۴۸	۷۸/۰۰	۱۹/۹۵	۱۰/۵۰	۶۲/۰۰	۹۹۰/۰۰	۸/۲۸	۴/۳۸	۱۰/۶۵	۰/۳۳	۴/۴۰
5H1	۷۴/۶۶	۲۳/۲۵	۱۸/۹۸	۱۱/۳۳	۱۰/۶۷	۱۱۳۰/۷۵	۷/۸۳	۸/۴۰	۲۳/۲۸	۰/۳۶	۱/۹۱
5A4	۶۴/۰۹	۱۵/۵۰	۱۹/۸۵	۱۱/۶۱	۸/۵۰	۱۰۹۸/۰۰	۸/۲۰	۵/۸۱	۱۹/۲۸	۰/۳۹	۱/۳۶
5H2	۲۸/۸۳	۴۴/۲۵	۲۲/۱۰	۹/۷۹	۱۴/۰۱	۱۱۶۲/۲۵	۸/۲۳	۶/۳۰	۱۸/۴۳	۰/۵۱	۱/۸۳
5P1	۶۱/۴۶	۴۶/۷۵	۲۲/۰۳	۸/۹۶	۱۷/۰۱	۱۴۶۲/۲۵	۸/۰۵	۶/۷۸	۱۶/۰۰	۰/۳۴	۱۰/۶۰
5A3	۶۳/۶۵	۸۴/۲۵	۲۳/۲۸	۹/۸۷	۱۴/۵۲	۱۷۶۰/۷۵	۸/۳۳	۴/۵۸	۱۹/۹۰	۰/۸۱	۲/۱۸
5A5	۰/۱۷	۱۸/۷۵	۲۳/۴۵	۱۰/۵۹	۲۵/۷۷	۷۹۷۴/۲۵	۸/۰۳	۶/۳۳	۳۳/۲۳	۰/۳۶	۵/۸۲
5A2	۵۵/۲۴	۴۷۶/۲۵	۲۲/۵۰	۸/۵۰	۳۵/۷۵	۲۴۱۶/۷۵	۸/۲۸	۴/۰۵	۱۷/۱۵	۰/۷۶	۳/۳۰
5C1	۵۶/۸۲	۸۷/۵۰	۲۱/۷۰	۹/۵۹	۸۰/۵۱	۲۰۵۱/۰۰	۸/۰۳	۵/۲۸	۲۲/۶۳	۰/۷۹	۵/۴۳
5A1	۳/۴۱	۵۰/۷۵	۲۲/۶۵	۹/۱۴	۹/۲۶	۲۷۱۵/۵۰	۸/۲۸	۸/۵۰	۱۸/۵۸	۱/۳۹	۵/۵۲
انحراف معیار	۲۶/۳۰	۱۵۷/۸۹	۳/۷۲	۱/۲۲	۱۶۵/۸۴	۱۷۱۳/۶۸	۰/۲۸	۲/۵۴	۶/۶۳	۰/۷۰	۴/۳۹

DO: Dissolved oxygen; TSS: Total suspended solids; EC: Electrical conductivity; BOD: Biological oxygen demand; COD: Chemical oxygen demand; TP: Total phosphor; TN: Total nitrogen

می‌شود، تفسیر مؤلفه‌ها مشکل خواهد بود. از این رو روش‌هایی به وجود آمده است که با رفع این مشکل، باعث تفسیر ساده‌تر مؤلفه‌ها شوند. این روش‌ها همان دوران مؤلفه‌ها هستند و به دو نوع دوران عمود و دوران مایل تقسیم می‌شوند. یکی از روش‌های چرخش عمودی که در مطالعات علمی بیشتر مورد استفاده قرار گرفته است، چرخش واریماکس (Varimax rotation) نامیده می‌شود (۲۳-۲۰). این روش نسبت به بقیه روش‌ها نتایج بهتری را در پی دارد و به عنوان چرخش استاندارد توصیه می‌گردد. این مرحله (استفاده از چرخش واریماکس جهت تفسیر بهتر نتایج) به PFA (Principal factor analysis) مشهور است (۱۹).

آنالیز همبستگی کانونیک (CCA)

در بسیاری از مطالعات آماری در عمل دسته‌ای از متغیرها را می‌توان به دو دسته متغیرهای پیش‌بینی کننده (مستقل) و متغیرهای پاسخ (وابسته) طبقه‌بندی نمود. در صورتی که تعداد متغیرهای پاسخ واحد باشد به راحتی می‌توان از روش‌هایی مانند رگرسیون خطی چند متغیره برای تجزیه و تحلیل روابط بین دو دسته متغیرها استفاده نمود، اما در برخی از موارد تعداد متغیرهای پاسخ بیشتر از یکی است و برای تجزیه و تحلیل روابط بین آن‌ها نمی‌توان از روش‌هایی مانند رگرسیون خطی چند متغیره استفاده نمود. در چنین مواردی CCA به عنوان گزینه مناسبی می‌تواند ارتباط بین این دو گروه از داده‌ها را

بررسی نماید. در این روش هدف رسیدن به دو دسته متغیر استاندارد مانند $U = aX$ و $V = bY$ است که ترکیبی خطی از متغیرهای اصلی می‌باشند (۱۰). در این روش مقادیر a و b طوری انتخاب می‌شوند که U و V بیشترین همبستگی را با یکدیگر داشته باشند. این روش تا به دست آوردن m مجموعه دسته از متغیرهای کانونیک ادامه پیدا می‌کند که m مینیمم تعداد متغیرهای گروه‌های مورد بررسی است (۲۴). در این پژوهش m بیانگر تعداد متغیرهای فیزیکی است. قابل ذکر است که بین U و V همبستگی خطی ساده‌ای وجود دارد. در واقع U و V از توزیع نرمال تبعیت کرده و واریانس آن‌ها واحد است. برای یافتن a و b باید ماتریس همبستگی بین متغیرهای X و Y تشکیل گردد. فرض کنید که ماتریس همبستگی $(p + q) \times (p + q)$ بین متغیرهای X_1 تا X_p مطابق شکل ۲ باشد:

$$\begin{array}{c|c} & \begin{array}{cc} X_1\ X_2 \ldots X_p & Y_1\ Y_2 \ldots Y_q \end{array} \\ \begin{array}{c} X_1 \\ X_2 \\ \vdots \\ X_p \\ Y_1 \\ Y_2 \\ \vdots \\ Y_q \end{array} & \left[\begin{array}{c:c} \begin{array}{c} p \times p \text{ matrix} \\ A \end{array} & \begin{array}{c} p \times q \text{ matrix} \\ C \end{array} \\ \hdashline \begin{array}{c} q \times p \text{ matrix} \\ C' \end{array} & \begin{array}{c} q \times q \text{ matrix} \\ B \end{array} \end{array}\right] \end{array}$$

شکل ۲: نحوه تشکیل ماتریس همبستگی بین متغیرها

این ماتریس از روی نمونه‌ای که در آن متغیرها اندازه‌گیری شده‌اند، محاسبه می‌گردد. از این ماتریس یک ماتریس $q \times q$ یعنی $B^{-1} C' A^{-1} C$ را می‌توان محاسبه نمود و مقدار ویژه آن را با حل معادله زیر محاسبه کرد:

فرمول ۳: $(\mathbf{B^{-1}C'A^{-1}C} - \lambda\mathbf{I})\mathbf{b} = 0$

در این حالت مقادیر ویژه $\lambda_1 > \lambda_2 > \lambda_3 > \ldots > \lambda_r$ عبارت از مربع مقادیر همبستگی بین دسته متغیرهای کانونیک می‌باشند. بردارهای ویژه b_1 تا b_r ضرایب متغیرهای Y را برای دسته متغیرهای کانونیک ارایه می‌دهد. ضریب U_i متغیر کانونیک **i** ام برای متغیرهای X، توسط معادله زیر به دست می‌آیند.

فرمول ۴: $\boldsymbol{\alpha}_i = \mathbf{A^{-1}Cb_i}$

جهت اطلاعات بیشتر در مورد تئوری روش CCA می‌توان به مراجع مربوط در این زمینه رجوع نمود (۲۵، ۲۴).

یافته‌ها

تعیین پارامترهای کیفی اصلی

برای تعیین پارمترهای کیفی اصلی آب، ۱۱ پارامتر برای این مطالعه انتخاب شد. اولین گام در تعیین پارامترهای کیفی اصلی با استفاده از روش آماری PCA، محاسبه ماتریس متقارن همبستگی θ می‌باشد (شکل ۳).

پس از حل معادله ۲، ۱۱ مقدار ویژه مطابق جدول ۲ محاسبه گردید. در این جدول، مقادیر ویژه و نسبت‌های واریانس تجمعی برای ترکیبات اصلی تولید شده نشان داده شده است. همان طور که ملاحظه می‌گردد، ۹۶/۹ درصد واریانس با نسبتی غیر قابل اغماض برای ۸ فاکتور اول تخصیص داده می‌شود.

جدول ۳ ضرایب تشکیل اجزا را نشان می‌دهد. در تحقیق حاضر، ضریب همبستگی اصلی قلمداد می‌شود که بزرگ‌تر از

$$\theta = \begin{bmatrix} 1.00 & 0.02 & 0.41 & -0.07 & -0.09 & -0.14 & 0.19 & -0.08 & -0.10 & -0.45 & -0.30 \\ 0.02 & 1.00 & 0.22 & 0.22 & 0.66 & -0.19 & 0.54 & -0.16 & -0.09 & -0.06 & 0.26 \\ 0.41 & -0.02 & 1.00 & -0.30 & -0.10 & 0.53 & 0.38 & -0.14 & -0.03 & -0.72 & -0.53 \\ -0.07 & 0.22 & -0.30 & 1.00 & 0.67 & -0.19 & 0.30 & 0.30 & 0.59 & 0.09 & 0.27 \\ -0.09 & 0.66 & -0.10 & 0.67 & 1.00 & -0.16 & 0.22 & 0.22 & 0.32 & 0.02 & 0.38 \\ -0.14 & -0.19 & 0.53 & -0.19 & -0.16 & 1.00 & -0.10 & -0.10 & 0.41 & -0.27 & -0.21 \\ 0.19 & 0.54 & 0.38 & 0.30 & 0.53 & -0.03 & 1.00 & 0.12 & -0.03 & -0.28 & 0.18 \\ -0.08 & -0.16 & -0.14 & 0.30 & 0.22 & -0.10 & 0.12 & 1.00 & 0.51 & 0.06 & 0.42 \\ -0.10 & -0.09 & -0.03 & 0.59 & 0.32 & 0.41 & -0.03 & 0.51 & 1.00 & 0.15 & 0.05 \\ -0.45 & -0.06 & -0.72 & 0.09 & 0.02 & -0.27 & -0.28 & 0.06 & 0.15 & 1.00 & 0.51 \\ -0.30 & 0.26 & -0.53 & 0.27 & 0.38 & -0.21 & 0.18 & 0.42 & 0.05 & 0.51 & 1.00 \end{bmatrix}$$

شکل ۳: ماتریس متقارن همبستگی θ

www.mui.ac.ir

جدول ۲: آماره‌های توصیفی مؤلفه‌های اصلی تولید شده

PCs*	مقادیر ویژه	واریانس نسبی	واریانس نسبی تجمعی
PC۱	۳/۲۰۲۳	۲۹/۱۱۲۳	۲۹/۱۱۲۳
PC۲	۲/۴۹۹۴	۲۲/۷۲۱۵	۵۱/۸۳۳۸
PC۳	۱/۶۳۱۳	۱۴/۸۳۰۴	۶۶/۶۶۴۲
PC۴	۱/۰۱۶۷	۹/۲۴۳۲	۷۵/۹۰۷۳
PC۵	۰/۸۵۵۱	۷/۷۷۳۴	۸۳/۶۸۰۷
PC۶	۰/۵۹۵۳	۵/۴۱۲۱	۸۹/۰۹۲۸
PC۷	۰/۴۹۸۲	۴/۵۲۹۳	۹۳/۶۲۲۱
PC۸	۰/۳۶۰۹	۳/۲۸۰۹	۹۶/۹۰۳۰
PC۹	۰/۱۸۳۰	۱/۶۶۳۲	۹۸/۵۶۶۲
PC۱۰	۰/۱۰۸۰	۰/۹۸۱۶	۹۹/۵۴۷۸
PC۱۱	۰/۰۴۹۷	۰/۴۵۲۲	۱۰۰

*PC: Principal component

۰/۵ (یا ۵۰ درصد) باشد. این معیار محافظه‌کارانه، به دلیل سطح مورد مطالعه بزرگ، عدم قطعیت در داده‌ها و سیستم غیر خطی و پویای رودخانه کرخه انتخاب می‌شود. در جدول ۳ مؤثرترین متغیرها در تشکیل اجزای اصلی به صورت پررنگ نشان داده شده‌اند. چنان که ملاحظه می‌گردد، تمام پارامترهای انتخاب شده طبق آنالیز آماری PCA مهم هستند. نتیجه‌گیری می‌شود که تمامی پارامترهای آب مورد استفاده در این تحقیق پارامترهای مهمی در نظر گرفته می‌شوند.

لازم به ذکر است که به طور قطع تمامی پارامترهای کیفی اندازه‌گیری شده در طول یک رودخانه نمی‌توانند پارامترهای مهمی از لحاظ کیفیت آن رودخانه باشند. این نتیجه تنها در مورد پارامترهای کیفی انتخاب شده در این تحقیق و تنها برای رودخانه کرخه معتبر می‌باشد. بنابراین انجام این آنالیز آماری برای دسته داده‌های مختلف برای رودخانه‌های مختلف می‌تواند نتایج مختلفی را در بر داشته و تنها تعدادی از پارامترهای کیفی انتخاب شده دارای ضرایب همبستگی اصلی بوده و مهم تلقی گردند.

رابطه بین پارامترهای فیزیکی و شیمیایی در حوضه رودخانه کرخه

در تحقیق حاضر، آنالیز CCA بر روی تمامی دسته اطلاعات موجود از حوضه آبریز انجام می‌شود. ۶ متغیر در سری داده‌های پاسخ وجود دارند، یعنی پارامترهای فیزیکی شامل Q، Turb، T، DO، TSS و EC و ۵ متغیر در سری داده‌های پیش‌بینی کننده یعنی پارامترهای شیمیایی شامل pH، BOD_5، COD، TP و TN. جدول ۴ نتایج CCA را برای متغیرهای فیزیکی و شیمیایی نشان می‌دهد.

قابل ذکر است که مقدار کوچک پارامتر سطح اهمیت (P)، بیانگر احتمال بالای رد فرضیه صفر مبنی بر بی‌ارتباط بودن دو دسته پارامترهای آلاینده فیزیکی و شیمیایی است (یعنی دارای ارتباط هستند). با توجه به نتایج به دست آمده در جدول ۴، بر مبنای نتایج سطح اهمیت (P) می‌توان دریافت که تنها ضریب همبستگی متغیر کانونیک اول دارای مقدار قابل قبول است و در واقع فرض صفر مبنی بر بی‌اهمیت بودن آن با احتمال ۸۵ درصد رد می‌شود (یعنی دارای اهمیت است). آزمون χ^2 نیز بیانگر عدم اهمیت ضرایب همبستگی متغیرهای کانونیک به جز مقدار ضریب همبستگی متغیرکانونیک اول است.

ضریب همبستگی برای متغیرهای استاندارد ۱، ۲ و ۳ به ترتیب ۰/۸۹۶، ۰/۸۴۸ و ۰/۵۹۳ بودند. ضریب همبستگی برای متغیر استاندارد پنجم و ششم کوچک‌تر از ۰/۵ می‌باشد و در نتیجه از آن‌ها صرف نظر شد. تنها همبستگی استاندارد اول از لحاظ آماری مهم بود ($P < ۰/۱۵۹$). آماره تست χ^2 برای متغیرهای استاندارد ۲ و ۳، به ترتیب $\chi^2_2 = ۱۹/۷۲$ با ۲۰ درجه آزادی و $\chi^2_3 = ۵/۷۱$ با ۱۲ درجه آزادی به دست آمد. اگر چه همبستگی‌های استاندارد ۲ و ۳ بزرگ بودند، بنابراین با استفاده از تست χ^2، از لحاظ آماری مهم قلمداد نمی‌شوند. بنابراین، مدرک مستدلی برای رابطه‌ای بین متغیرهای فیزیکی و شیمیایی بر مبنای این متغیرهای استاندارد وجود ندارد. ممکن است عجیب به نظر برسد که اگر چه این همبستگی‌های استاندارد به نسبت بالا هستند، هیچ درجه اهمیتی در نتایج مشاهده نمی‌شود. اندازه به نسبت کوچک نمونه‌های تهیه شده شاید توضیح منطقی برای این موضوع باشد (۲۶).

متغیرهای غالب در اولین متغیر استاندارد در جدول ۴ (که از لحاظ آماری مهم قلمداد می‌شود)، برای متغیرهای فیزیکی U_1، دبی (Q) و دما (T) بوده و متغیرهای غالب در V_1 (پارامترهای شیمیایی)، pH، TP و TN هستند. بقیه

www.mui.ac.ir

جدول ۳: ضرایب ترکیبات اصلی تولید شده

	$PC_{۱}$	$PC_{۲}$	$PC_{۳}$	$PC_{۴}$	$PC_{۵}$	$PC_{۶}$	$PC_{۷}$	$PC_{۸}$	$PC_{۹}$	$PC_{۱۰}$	$PC_{۱۱}$
Q	۰/۳۴۹	−۰/۴۶۵	۰/۳۰۰	۰/۶۰۰	−۰/۰۰۴	۰/۲۲۳	۰/۳۸	۰/۱۳۰	−۰/۰۱۵	۰/۰۳۹	−۰/۰۴۱
Turb.	−۰/۵۲۴	−۰/۵۲۱	−۰/۲۷۱	−۰/۲۸۸	−۰/۰۷۹	۰/۵۲۳	−۰/۱۷۹	۰/۰۷۰	−۰/۱۴۷	۰/۰۹۰	۰/۰۳۶
T	۰/۶۱۴	−۰/۷۰۰	۰/۱۶۴	−۰/۰۹۹	−۰/۱۲۹	−۰/۰۷۲	۰/۰۳۸	−۰/۱۱۶	۰/۱۸۵	۰/۱۱۸	۰/۱۱۹
DO	−۰/۶۸۶	−۰/۲۵۹	۰/۱۲۵	۰/۲۶۴	۰/۵۱۴	−۰/۳۴۶	−۰/۰۰۲	۰/۰۲۲	−۰/۱۳۵	۰/۱۲۷	۰/۰۴۵
TSS	−۰/۷۴۲	−۰/۵۰۸	−۰/۰۳۷	−۰/۱۴۵	۰/۲۵۷	۰/۰۲۴	−۰/۰۳۲	۰/۱۱۸	۰/۲۸۶	۰/۰۱۹	−۰/۰۸۳
EC	۰/۳۷۴	−۰/۲۲۲	۰/۷۵۷	−۰/۳۹۳	−۰/۰۵	−۰/۰۴۱	۰/۱۷۳	۰/۱۵۶	−۰/۱۰۱	۰/۰۹۰	−۰/۰۹۳
pH	−۰/۲۸۸	−۰/۷۶۶	−۰/۲۴۱	−۰/۲۱۲	−۰/۱۴۷	−۰/۱۹۵	۰/۲۰۵	−۰/۳۲۰	۰/۱۱۰	−۰/۱۰۹	−۰/۰۳۹
BOD5	−۰/۵۲۴	−۰/۱۰۵	۰/۳۵۳	۰/۴۵۰	−۰/۵۵۷	−۰/۰۴۸	−۰/۲۳۶	−۰/۱۱۳	۰/۰۲۵	۰/۰۷۷	−۰/۰۵۸
COD	−۰/۳۶۳	−۰/۱۶۳	۰/۸۱۵	۰/۱۹۶	۰/۱۳۷	۰/۲۸۵	۰/۰۹۳	−۰/۰۵۸	−۰/۰۰۷	−۰/۱۵۵	۰/۰۷۰
TP	−۰/۴۸۹	۰/۶۹۹	۰/۰۲۴	−۰/۱۲۸	۰/۰۶۶	۰/۲۰۲	۰/۳۱۸	−۰/۲۹۶	۰/۰۶۹	۰/۱۳۵	−۰/۰۱۱
TN	−۰/۷۳۸	۰/۱۹۶	−۰/۰۸۲	−۰/۱۴۱	−۰/۴۳۳	−۰/۱۷۹	۰/۲۸۹	۰/۲۸۶	۰/۰۲۱	−۰/۰۲۶	۰/۰۷۷

DO: Dissolved oxygen; TSS: Total suspended solids; EC: Electrical conductivity; BOD: Biological oxygen demand; COD: Chemical oxygen demand; TP: Total phosphor; TN: Total nitrogen

جدول ۴: آنالیز همبستگی کانونیک داده‌های کیفی

متغیرهای همبستگی	همبستگی متعارف	$\chi^{۲}$	درجه آزادی	سطح اهمیت (P)	پارامترهای فیزیکی						پارامترهای شیمیایی				
					Q	Turb	T	DO	TSS	EC	pH	BOD_5	COD	TP	TN
۱	۰/۸۹۶	۳۷/۶۳۵	۳۰	< ۰/۱۵۹	−۰/۶۶۲	−۰/۳۵۶	−۰/۸۴۴	−۰/۱۷۴	−۰/۳۱۹	۰/۳۰۰	−۰/۷۲۳	−۰/۰۲۴	−۰/۱۱۲	۰/۷۸۴	۰/۵۳۲
۲	۰/۸۴۸	۱۹/۷۲۳	۲۰	< ۰/۴۷۵	۰/۱۹۵	−۰/۲۳۷	۰/۱۸۵	−۰/۶۰۱	−۰/۶۲۱	−۰/۵۶۴	−۰/۰۸۰	−۰/۱۶۲	−۰/۷۸۱	−۰/۲۰	−۰/۴۲۴
۳	۰/۵۹۳	۵/۷۰۵	۱۲	< ۰/۹۳۰	−۰/۲۴۴	۰/۷۲۵	−۰/۴۱۴	۰/۴۵۷	۰/۷۱۲	−۰/۶۶۸	۰/۶۶۳	۰/۳۷۹	−۰/۱۵۱	۰/۳۷۳	۰/۶۸۱
۴	۰/۲۷۷	۰/۹۳۰	۶	< ۰/۹۸۸	۰/۱۹۴	۰/۲۲۵	۰/۲۰۸	−۰/۶۱۸	۰/۱۱۱	۰/۰۳۳	۰/۱۳۷	−۰/۷۱۰	−۰/۵۹۳	−۰/۱۹۵	۰/۱۸۳
۵	۰/۰۷۰۱	۰/۹۷۴	۲	< ۰/۹۷۳	۰/۸۰۵	۰/۰۲۱	−۰/۱۹۴	۰/۲۲۹	۰/۰۳۶	۰/۴۵۰	−۰/۱۱۲	۰/۵۷۱	−۰/۰۵۷	−۰/۴۱۰	۰/۴۰۱

DO: Dissolved oxygen; TSS: Total suspended solids; EC: Electrical conductivity; BOD: Biological oxygen demand; COD: Chemical oxygen demand; TP: Total phosphor; TN: Total nitrogen

www.mui.ac.ir

پارامترهای فیزیکی و شیمیایی در این متغیر استاندارد روابط مهمی را نشان ندادند. در متغیر استاندارد دوم، متغیرهای پیش‌بینی کننده، DO و TSS هستند و از بین متغیرهای پاسخ نیز EC و COD دارای همبستگی بالایی می‌باشد. از همبستگی استاندارد سوم نیز، Turb، TSS و EC از بین دسته متغیرهای U_1، pH و TN از بین دسته متغیرهای V_1 و در همبستگی استاندارد چهارم تنها DO از بین متغیرهای فیزیکی و BOD_5 و COD از بین پارامترهای شیمیایی دارای همبستگی بالایی می‌باشد. در همبستگی استاندارد پنجم هیچ ارتباط معنی‌داری بین پارامترهای فیزیکی و شیمیایی وجود ندارد. متغیر استاندارد پنجم نشان می‌دهد که غیر از ارتباط معنی‌دار BOD_5 با Q، هیچ ارتباط معنی‌دار دیگری بین بقیه پارامترها وجود ندارد.

با بررسی قابلیت CCA، فاکتور همبستگی ساده بین تمامی پارامترها در جدول ۵ نشان داده شده است. ماتریس همبستگی معلوم می‌کند که یک رابطه EC و T، بین TSS و Turb و بین DO و TSS از بین پارامترهای فیزیکی و همچنین بین COD و BOD_5 و بین TP و TN از بین پارامترهای شیمیایی وجود دارد، همچنین رابطه معنی‌داری بین پارامتر فیزیکی T با پارامترهای شیمیایی TP و TN و بین DO با COD وجود دارد.

بحث

داده‌های کیفی آب سطحی برای رودخانه کرخه (در غرب ایران) برای بررسی تغییرات مکانی و رابطه بین پارامترهای فیزیکی و شیمیایی ارزیابی شد. آنالیز PCA برای تعیین پارامترهای کیفی اصلی آب و آنالیز CCA برای یافتن رابطه بین دو سری داده‌ها، یعنی پارامترهای فیزیکی و شیمیایی به کار برده شد. نتایج PCA نشان داد که همه پارامترهای کیفی انتخاب شده برای رودخانه کرخه در این تحقیق، اصلی و مهم هستند. لازم به ذکر است که به طور قطع تمام پارامترهای اندازه‌گیری شده در طول یک رودخانه نمی‌توانند پارامترهای اصلی مسؤول کیفیت رودخانه باشند. این نتیجه تنها برای رودخانه کرخه و پارامترهای انتخاب شده در این تحقیق معتبر می‌باشد. در عین حال بسته به منشأ پارامترهای کیفی در حوضه آبریز هر رودخانه، نتایج آنالیزهای آماری ممکن است منجر به انتخاب تعداد خاصی از پارامترها (و نه تمامی پارامترهای کیفی) به عنوان پارامترهای کیفی اصلی در حوضه آبریز رودخانه گردد. در تأیید این مطلب می‌توان به تحقیق Ouyang (۱۷) اشاره نمود. وی PCA را برای شناسایی

جدول ۵: ماتریس همبستگی داده‌های کیفی

	Q	Turb.	T	DO	TSS	EC	pH	BOD5	COD	TP	TN
Q	۱/۰۰										
Turb.	۰/۰۲	۱/۰۰									
T	۰/۴۱	−۰/۰۲	۱/۰۰								
DO	−۰/۰۷	۰/۲۲	−۰/۳۰	۱/۰۰							
TSS	−۰/۰۹	۰/۶۶	−۰/۱۰	۰/۶۷	۱/۰۰						
EC	−۰/۱۴	−۰/۱۹	۰/۵۳	−۰/۱۹	−۰/۱۶	۱/۰۰					
pH	۰/۱۹	۰/۵۴	۰/۳۸	۰/۳۰	۰/۵۳	−۰/۰۳	۱/۰۰				
BOD5	−۰/۰۸	−۰/۱۶	−۰/۱۴	۰/۳۰	۰/۲۲	−۰/۱۰	۰/۱۲	۱/۰۰			
COD	−۰/۱۰	−۰/۰۹	−۰/۰۳	۰/۵۹	۰/۳۲	۰/۴۱	−۰/۰۳	۰/۵۱	۱/۰۰		
TP	−۰/۴۵	−۰/۰۶	−۰/۷۲	۰/۰۹	۰/۰۲	−۰/۲۷	−۰/۲۸	۰/۰۶	۰/۱۵	۱/۰۰	
TN	−۰/۳۰	۰/۲۶	−۰/۵۳	۰/۲۷	۰/۳۸	−۰/۲۱	۰/۱۸	۰/۴۲	۰/۰۵	۰/۵۱	۱/۰۰

DO: Dissolved oxygen; TSS: Total suspended solids; EC: Electrical conductivity;BOD: Biological oxygen demand
COD: Chemical oxygen demand; TP: Total phosphor; TN: Total nitrogen

www.mui.ac.ir

پارامترهای کیفی مهم آب در ۲۲ ایستگاه واقع در بخش اصلی رودخانه St.Johns سفلی در فلوریدای آمریکا به کار برد. نتایج نشانگر این بودند که از بین تمامی پارامترهای مورد بررسی در این تحقیق، تنها کل کربن آلی TOC، کربن آلی محلول DOC، نیتروژن کل TN، نیترات و نیتریت محلول، اورتوفسفات، قلیائیت، شوری، منیزیم و کلسیم، مهم‌ترین پارامترها در ارزیابی متغیرهای کیفی آب رودخانه بودند.

در تحقیق دیگری که توسط نوری و همکاران (۲۶) صورت گرفت، آنالیز PCA برای تعیین ایستگاه‌های کیفی اصلی و غیر اصلی از بین ۱۷ ایستگاه اندازه‌گیری کیفی واقع در طول رودخانه کارون را مورد استفاده قرار دادند. آن‌ها همچنین از این آنالیز آماری جهت استخراج پارامترهای کیفی اصلی مسؤول مشکلات کیفی رودخانه کارون از بین پارامترهای COD، BOD_5، EC، NO_3^-، SO_4^{2-}، pH، TDS، سختی (Hardness)، DO، Cl^-، T و کدورت (Turbidity) اندازه‌گیری شده در ایستگاه‌های اصلی تعیین شده در مرحله اول تحقیقشان استفاده کردند. آن‌ها به این نتیجه رسیدند که تمام پارامترهای کیفی مورد بررسی، اصلی می‌باشد و مسؤول اصلی مشکلات کیفی رودخانه هستند.

بر اساس آنالیز آماری CCA برای پارامترهای کیفی رودخانه کرخه در تحقیق حاضر، Turb، DO و TSS پارامترهای فیزیکی مهم و BOD_5، COD، TP و TN پارامترهای شیمیایی مهم در متغیرهای کانونیک بودند.

نتایج به دست آمده از این تحقیق می‌تواند بیانگر ارتباط معنی‌داری بین دو دسته پارامترهای آلاینده شیمیایی و فیزیکی و در نتیجه منشأ یکسان آن‌ها در رودخانه کرخه باشد. به عبارتی می‌توان ابراز داشت که هر دو دسته پارامترهای آلاینده شیمیایی و فیزیکی از یک منشأ سرچشمه می‌گیرند. لازم به ذکر است که در حوضه آبریز یک رودخانه، آلاینده‌های شیمیایی مانند نیتروژن و فسفر بیشتر ناشی از فعالیت‌های انسان‌ساخت مانند پساب‌های شهری، کشاورزی و صنعتی می‌باشد، اما پارامترهای آلاینده فیزیکی می‌توانند هم منشأ انسان‌ساخت و هم طبیعی (ناشی از خصوصیات فیزیکی حوضه مانند زمین‌شناسی و خاک‌شناسی) داشته باشند. بنابراین بر اساس نتایج به دست آمده از تحلیل همبستگی کانونیک برای داده‌های کیفی مورد بررسی در این تحقیق، مشخص می‌شود که منشأ هر دو دسته پارامترهای آلاینده شیمیایی و فیزیکی، انسان‌ساخت می‌باشد.

به عنوان جمع‌بندی نتایج این تحقیق، می‌توان ابراز داشت که روش‌های مورد استفاده با اکتشاف پارامترهای کیفی اصلی و ارتباط بین پارامترهای کیفی فیزیکی و شیمیایی رودخانه کرخه می‌تواند به عنوان راه‌حل مؤثری در مدیریت کیفی آب در رودخانه کرخه چنان‌که داده‌های کیفی بزرگ مقیاس و پیچیده آب در دسترس باشد، به کار رود.

References

1. Noori P, Abdoli MA, Ameri Ghasrodashti, Ghazizade M. Prediction of municipal solid waste generation with combination of support vector machine and principal component analysis: A case study of Mashhad. Environmental Progress & Sustainable Energy 2009; 28(2): 249-58.
2. Noori R, Karbassi AR, Farokhnia A, Dehghani M. Predicting the Longitudinal Dispersion Coefficient Using Support Vector Machine and Adaptive Neuro-Fuzzy Inference System Techniques. Environmental Engineering Science 2009; 26(10): 1503-10
3. Shrestha S, Kazama F. Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. Environmental Modelling & Software 2007; 22(4): 464-75.
4. Dixon W, Chiswell B. Review of aquatic monitoring program design. Water Research 1996; 30(9): 1935-48.
5. Nouri R, Kerachian R, Khodadadi Darban A, Shakibaei Nia A. Assessment of importance of water quality monitoring stations using principal components analysis and factor analysis: a case study of the Karoon river. Water and Wastewater 2007; 18(3): 60-9. [In Persian].
6. Barnett TP, Preisendorfer R. Origins and Levels of Monthly and Seasonal Forecast Skill for United States Surface Air Temperatures Determined by Canonical Correlation Analysis. Journal: Monthly Weather Review 1987; 115(9): 1825-50.
7. Statheropoulos M, Vassiliadis N, Pappa A. Principal component and canonical correlation analysis for examining air pollution and meteorological data. Atmospheric Environment 1998; 32(6): 1087-95.

www.mui.ac.ir

8. Larson M, Capobianco M, Hanson H. Relationship between beach profiles and waves at Duck, North Carolina, determined by canonical correlation analysis. Marine Geology 2000; 163(1): 275-88.
9. Ouarda TBMJ, Girard C, Cavadias GS, Bobee B. Regional flood frequency estimation with canonical correlation analysis. Journal of Hydrology 2001; 254(1-4): 157-73.
10. Gramotnev G, Gramotnev DK. Multi-channel statistical analysis of combustion aerosols. Part I: Canonical correlations and sources of particle modes. Atmospheric Environment 2007; 41(17): 3521-34.
11. Rasuli A, Noori R, Vesali-Naseh MR, Nazariha M, Kiaghadi A. Determination of the Relationship Between Physical and Chemical Pollutant Parameters in the Sefidrood River Basin Based on Canonical Correlation Analysis. Proceedings of 4th Iran Water Resources Management Conference; 2011 May 23- 24; Tehran, Iran; 2011. [In Persian].
12. Terrado M, Barcelo D, Tauler R. Identification and distribution of contamination sources in the Ebro river basin by chemometrics modelling coupled to geographical information systems. Talanta 2006; 70(4): 691-704.
13. Simeonov V, Stratis JA, Samara C, Zachariadis G, Voutsa D, Anthemidis A, et al. Assessment of the surface water quality in Northern Greece. Water Res 2003; 37(17): 4119-24.
14. Singh KP, Malik A, Mohan D, Sinha S. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)-a case study. Water Res 2004; 38(18): 3980-92.
15. Kalin M, Cao Y, Smith M, Olaveson MM. Development of the phytoplankton community in a pit-lake in relation to water quality changes. Water Res 2001; 35(13): 3215-25.
16. Gangopadhyay S, Gupta A, Nachabe MH. Evaluation of ground water monitoring network by principal component analysis. Ground Water 2001; 39(2): 181-91.
17. Ouyang Y. Evaluation of river water quality monitoring stations by principal component analysis. Water Res 2005; 39(12): 2621-35.
18. Noori R, Farokhnia A, Riahi Medvar H. Effects of preprocessing the inlet parameters to artificial neural network (ANN) for predicting the monthly flow by PCA and wavelet. Journal of Water & Wastewater 2009; 20(69): 13-22. [In Persian].
19. 19-. Manly BF. Multivariate Statistical Methods: A Primer. London, UK: Chapman and Hall; 1986
20. Morales MM, Marti P, Llopis A, Campos L, Sagrado S. An environmental study by factor analysis of surface seawaters in the Gulf of Valencia (Western Mediterranean). Analytica Chimica Acta 1999; 394(1): 109-17
21. Helena B, Pardo R, Vega M, Barrado E, Fernandez JM, Fernandez L. Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. Water Research 2000; 34(3): 807-16.
22. Noori R, Ashrafi Kh, Ajdarpour A. Comparison of ANN and PCA based multivariate linear regression applied to predict the daily average concentration of CO: a case study of Tehran. Journal of the Earth and Space Physics 2008; 34(2): 135-52.
23. Noori R, Sabahi MS, Karbassi AR, Baghvand A, Taati Zadeh H. Multivariate statistical analysis of surface water quality based on correlations and variations in the data set. Desalination 2010; 260(1-3): 129-36.
24. Graham NE, Michaelsen J, Barnett TP. An investigation of the El Nino-Southern Oscillation cycle with statistical models: 1. Predictor field characteristics. Journal of Geophysical Research 1987; 92(13): 14251-70.
25. Shabbar A, Barnston AG. Skill of Seasonal Climate Forecasts in Canada Using Canonical Correlation Analysis. Monthly Weather Review 1996; 124(10): 2370-85.
26. Diagomanolin V, Farhang M, Ghazi-Khansari M, Jafarzadeh N. Heavy metals (Ni, Cr, Cu) in the Karoon waterway river, Iran. Toxicology Letters 2004; 151(1): 63-7.

www.mui.ac.ir

Karkheh River Water Quality Using Multivariate Statistical Analysis and Qualitative Data Variations

Saeed Rasi Nezami[1], Mehrdad Nazariha[2], Akbar Baghvand[3], Ali Moridi[4]

Original Article

Abstract

Background: The spread of urbanization in floodplains as well as increase of industrial and agricultural activities in river discharges may lead to water pollution by physical, chemical and biological pollutants. One of the major concerns in monitoring the water resources system is indentifying major quality parameters and determining a relationship between physical and chemical parameters. This is important because the physical parameters are usually related to the natural characteristics of the river basin, while chemical parameters are associated with human-made pollution resources. Thus, if there is a high correlation between these two groups of parameters, it could be concluded that the source of both parameters are identical. However, to achieve this goal, costly and time consuming instrumental techniques and intensive laboratory efforts might be required.

Methods: To determine the major quality parameters that cause water pollution in Karkheh River, statistical principal components analysis (PCA) method was employed. Moreover, canonical correlation analysis (CCA) was used to find the relationships between physical and chemical parameters. In this study, discharge rate (Q), turbidity (Turb), temperature (T), dissolved oxygen (DO), total suspended solids (TSS) and electrical conductivity (EC) were selected as representative parameters for the physical characteristics of the river water quality. In addition, pH, 5-day biological oxygen demand (BOD_5), chemical oxygen demand (COD), total phosphorous (TP) and total nitrogen (TN) were chosen as representatives of the river chemical attributes. These two groups of quality parameters were measured along the Karkheh River in 18 monitoring stations from March 2006 to March 2007.

Findings: Results of PCA analysis indicated that all the selected chemical and physical parameters of Karkheh River were of high importance. Further investigations indicated that in CCA analysis the first three canonical correlations were 0.896, 0.848 and 0.593, respectively, suggesting that turbidity, DO and TSS were three dominant physical parameters in all the canonical variables, whilst BOD_5, COD, TP and TN were highly scored among the chemical parameters.

Conclusion: According to the CCA results, it can be implied that human activities are the source of the physical and chemical parameters of the river. The utilized statistical methods in this study can be used as an appropriate decision support system for water sector managers regarding water quality management of the river basin when large complex water quality data are involved.

Key words: Karkheh River, Principle Component Analysis, Canonical Correlation Analysis, Multivariate Statistical Analysis

Citation: Rasi Nezami S, Nazariha M, Baghvand A, Moridi A. **Karkheh River Water Quality Using Multivariate Statistical Analysis and Qualitative Data Variations.** J Health Syst Res 2013; 8(7): 1280-92.

Received date: 12/05/2012 Accept date: 16/08/2012

1- PhD Candidate, Civil and Environmental Engineering, Department of Environmental Engineering, University of Tehran, Tehran, Iran (Corresponding Author) Email: saeedrasinezami@yahoo.com
2- Assistant Professor, Department of Environmental Engineering, School of Environmental, University of Tehran, Tehran, Iran
3- Associate Professor, Department of Environmental Engineering, School of Environmental, University of Tehran, Tehran, Iran
4- Assistant Professor, Department of Water Resources, School of Water and Environmental Engineering, Water and Power Industry Shahid Abbaspour University, Tehran, Iran